HZ BOOKS 华章教育

McGraw Hill Education

国外电子与电气工程技术丛书

基于运算放大器的模拟集成电路设计

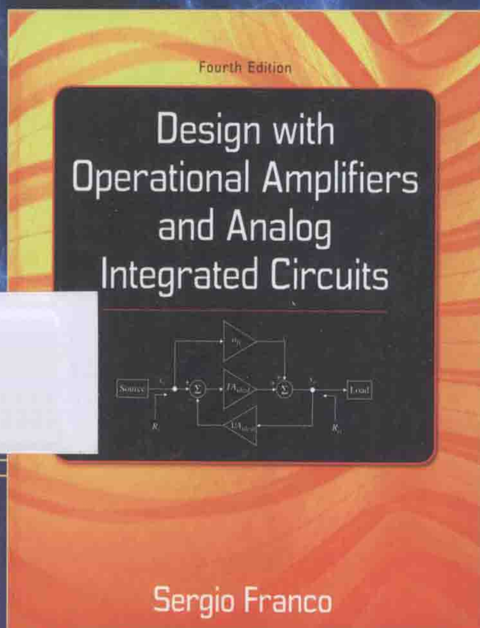
（英文版·第4版）

[美] Sergio Franco 著

何乐年 注释

Design with Operational Amplifiers and Analog Integrated Circuits

Fourth Edition



机械工业出版社
China Machine Press



本书注重理论和实际应用相结合，重点阐述模拟集成电路设计的原理和分析方
运算放大器的基本原理和应用、运算放大器的静态和动态限制、噪声及稳定性问题等诸多实际问题，强调物理思想，帮助读者提高电路设计的能力。

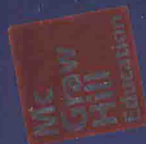
本书通俗易懂，可作为电子信息、通信、控制、仪器仪表等相关专业本科生及研究生的教材或主要参考书，对电子工程师也是一本实用的参考书。

与前一版相比，本书有如下特色：

- 更新了负反馈电路部分，涵盖了开关稳压器、锁相环（PLL）的反馈电路，并包括了双端口分析和回馈比例分析方法。
- 增加了运算放大器的动态特性和频率补偿分析内容，在环路增益测量方面，采用了由R. D. Middlebrook提出的电压/电流注入技术方法。
- 拓展了开关稳压器的内容，重点讨论和介绍了右半平面零点和误差放大器稳定性，以及电流控制模式和斜坡补偿技术。
- 增强了PSpice的使用，采用电路结构图进行验证计算，并且分析更高阶的效应。
- 重新设计了例题并且更新了章后习题。

作者简介

Sergio Franco 出生在意大利，1980年开始在美国旧金山州立大学电气工程系授课，期间获得了伊利诺伊大学香槟分校博士学位，成为该系荣誉教授。在就任现职之前，Franco博士拥有广泛的行业经验，在诸如固态物理学、模式识别、集成电路（IC）设计、医学电子、日用电子和汽车电子等领域工作过，发表论文颇丰。Franco博士还是《Analog Circuit Design: Discrete & Integrated》（McGraw-Hill Education，2014）和《Electric Ciruit Fundamentals》（Oxford University Press，1995）两本教科书的作者。





上架指导：模拟集成电路设计

ISBN 978-7-111-48933-7



定价：99.00元



www.mheducation.com

投稿热线：(010) 88379604
客服热线：(010) 88378991 88361066
购书热线：(010) 68326294 88379649 68995259

华章网站：www.hzbook.com
网上购书：www.china-pub.com
数字阅读：www.hzmedia.com.cn

封面设计 · 锡 彬

国外电子与电气工程技术丛书

基于运算放大器的模拟集成电路设计

（英文版·第4版）

[美] Sergio Franco 著
何乐年 注释



机械工业出版社
China Machine Press

图书在版编目（CIP）数据

基于运算放大器的模拟集成电路设计（英文版·第4版）/（美）弗兰克（Franco, S.）著；何乐年注释．—北京：机械工业出版社，2015.1
（国外电子与电气工程技术丛书）
书名原文：Design with Operational Amplifiers and Analog Integrated Circuits, Fourth Edition

ISBN 978-7-111-48933-7

I. 基… II. ① 弗… ② 何… III. ① 模拟集成电路－电路设计－英文 ② 运算放大器－电路设计－英文 IV. ① TN722.702 ② TN431.102

中国版本图书馆CIP数据核字（2014）第303199号

本书版权登记号：图字：01-2014-7782

出版发行：机械工业出版社（北京市西城区百万庄大街22号 邮政编码：100037）
责任编辑：谢晓芳　　责任校对：殷　虹
印　　刷：北京诚信伟业印刷有限公司　　版　　次：2015年1月第1版第1次印刷
开　　本：186mm×240mm 1/16　　印　　张：45.5
书　　号：ISBN 978-7-111-48933-7　　定　　价：99.00元

凡购本书，如有缺页、倒页、脱页，由本社发行部调换
客服热线：(010) 88378991 88361066　　投稿热线：(010) 88379604
购书热线：(010) 68326294 88379649 68995259　　读者信箱：hzjsj@hzbook.com

出版者的话

文艺复兴以降，源远流长的科学精神和逐步形成的学术规范，使西方国家在自然科学的各个领域取得了垄断性的优势；也正是这样的传统，使美国在信息技术发展的六十多年间名家辈出、独领风骚。在商业化的进程中，美国的产业界与教育界越来越紧密地结合，信息学科中的许多泰山北斗同时身处科研和教学的最前线，由此而产生的经典科学著作，不仅擘划了研究的范畴，还揭示了学术的源变，既遵循学术规范，又自有学者个性，其价值并不会因年月的流逝而减退。

近年，在全球信息化大潮的推动下，我国的信息产业发展迅猛，对专业人才的需求日益迫切。这对我国教育界和出版界都既是机遇，也是挑战；而专业教材的建设在教育战略上显得举足轻重。在我国信息技术发展时间较短的现状下，美国等发达国家在其信息科学发展的几十年间积淀和发展的经典教材仍有许多值得借鉴之处。因此，引进一批国外优秀教材将对我国教育事业的发展起到积极的推动作用，也是与世界接轨、建设真正的世界一流大学的必由之路。

机械工业出版社华章公司较早意识到“出版要为教育服务”。自 1998 年开始，我们就将工作重点放在了遴选、移译国外优秀教材上。经过多年的不懈努力，我们与 Pearson、McGraw-Hill、Elsevier、John Wiley & Sons、CRC、Springer 等世界著名出版公司建立了良好的合作关系，从他们现有的数百种教材中甄选出 Thomas L. Floyd、Charles K. Alexander、Behzad Razavi、John G. Proakis、Stephen Brown、Allan R. Hambley、Albert Malvino、Mark I.Montrose、David A. Johns 、Peter Wilson 、H. Vincent Poor、Dikshitulu K. Kalluri、Bhag Singh Guru、Stephane Mallat 等大师名家的经典教材，以“国外电子与电气工程技术丛书”为总称出版，供读者学习、研究及珍藏。这些书籍在读者中树立了良好的口碑，并被许多高校采用为正式教材和参考书籍。其影印版“经典原版书库”作为姊妹篇也越来越多被实施双语教学的学校所采用。

权威的作者、经典的教材、一流的译者、严格的审校、精细的编辑，这些因素使我们的图书有了质量的保证。随着电气与电子信息学科建设的不断完善和教材改革的逐渐深化，教育界对国外电气与电子信息教材的需求和应用都将步入一个新的阶段，我们的目标是尽善尽美，而反馈的意见正是我们达到这一终极目标的重要帮助。华章公司欢迎老师和读者对我们的工作提出建议或给予指正，我们的联系方法如下：

华章网站：www.hzbook.com

电子邮件：hzjsj@hzbook.com

联系电话：(010) 88379604

联系地址：北京市西城区百万庄南街 1 号

邮政编码：100037



华章科技图书出版中心

PREFACE

During the last decades much has been prophesized that there will be little need for analog circuitry in the future because digital electronics is taking over. Far from having proven true, this contention has provoked controversial rebuttals, as epitomized by statements such as "If you cannot do it in digital, it's got to be done in analog." Add to this the common misconception that analog design, compared to digital design, seems to be more of a whimsical art than a systematic science, and what is the confused student to make of this controversy? Is it worth pursuing some coursework in analog electronics, or is it better to focus just on digital?

There is no doubt that many functions that were traditionally the domain of analog electronics are nowadays implemented in digital form, a popular example being offered by digital audio. Here, the analog signals produced by microphones and other acoustic transducers are suitably conditioned by means of amplifiers and filters, and are then converted to digital form for further processing, such as mixing, editing, and the creation of special effects, as well as for the more mundane but no less important tasks of transmission, storage, and retrieval. Finally, digital information is converted back to analog signals for playing through loudspeakers. One of the main reasons why it is desirable to perform as many functions as possible digitally is the generally superior reliability and flexibility of digital circuitry. However, *the physical world is inherently analog,* indicating that there will *always* be a need for analog circuitry to condition physical signals such as those associated with transducers, as well as to convert information from analog to digital for processing, and from digital back to analog for reuse in the physical world. Moreover, new applications continue to emerge, where considerations of speed and power make it more advantageous to use analog front ends; wireless communications provide a good example.

Indeed many applications today are best addressed by mixed-mode integrated circuits (mixed-mode ICs) and systems, which rely on analog circuitry to interface with the physical world, and digital circuitry for processing and control. Even though the analog circuitry may constitute only a small portion of the total chip area, it is often the most challenging part to design as well as the limiting factor on the performance of the entire system. In this respect, it is usually the analog designer who is called to devise ingenious solutions to the task of realizing analog functions in decidedly digital technologies; switched-capacitor techniques in filtering and sigma-delta techniques in data conversion are popular examples. In light of the above, the need for competent analog designers will continue to remain very strong. Even purely digital circuits, when pushed to their operational limits, exhibit analog behavior. Consequently, a solid grasp of analog design principles and techniques is a valuable asset in the design of any IC, not just purely digital or purely analog ICs.

THE BOOK

The goal of this book is the illustration of general analog principles and design methodologies using practical devices and applications. The book is intended as a

textbook for undergraduate and graduate courses in design and applications with analog integrated circuits (analog ICs), as well as a reference book for practicing engineers. The reader is expected to have had an introductory course in electronics, to be conversant in frequency-domain analysis techniques, and to possess basic skills in the use of SPICE. Though the book contains enough material for a two-semester course, it can also serve as the basis for a one-semester course after suitable selection of topics. The selection process is facilitated by the fact that the book as well as its individual chapters have generally been designed to proceed from the elementary to the complex.

At San Francisco State University we have been using the book for a sequence of two one-semester courses, one at the senior and the other at the graduate level. In the senior course we cover Chapters 1–3, Chapters 5 and 6, and most of Chapters 9 and 10; in the graduate course we cover all the rest. The senior course is taken concurrently with a course in analog IC fabrication and design. For an effective utilization of analog ICs, it is important that the user be cognizant of their internal workings, at least qualitatively. To serve this need, the book provides intuitive explanations of the technological and circuital factors intervening in a design decision.

NEW TO THE FOURTH EDITION

The key features of the new edition are: *(a)* a complete revision of negative feedback, *(b)* much enhanced treatment of op amp dynamics and frequency compensation, *(c)* expanded coverage of switching regulators, *(d)* a more balanced presentation of bipolar and CMOS technologies, *(e)* a substantial increase of in-text PSpice usage, and *(f)* redesigned examples and about 25% new end-of-chapter problems to reflect the revisions.

While previous editions addressed negative feedback from the specialized viewpoint of the op amp user, the fourth edition offers a much broader perspective that will prove useful also in other areas like switching regulators and phase-locked loops. The new edition presents both two-port analysis and return-ratio analysis, emphasizing similarities but also differences, in an attempt at dispelling the persisting confusion between the two (to keep the distinction, the loop gain and the feedback factor are denoted as L and b in two-port analysis, and as T and β in return-ratio analysis).

Of necessity, the feedback revision is accompanied by an extensive rewriting of op amp dynamics and frequency compensation. In this connection, the fourth edition makes generous use of the voltage/current injection techniques pioneered by R. D. Middlebrook for loop-gain measurements.

In view of the importance of portable-power management in today's analog electronics, this edition offers an expanded coverage of switching regulators. Much greater attention is devoted to current control and slope compensation, along with stability issues such as the effect of the right-half plane zero and error-amplifier design.

The book makes abundant use of SPICE (schematic capture instead of the netlists of the previous editions), both to verify calculations and to investigate higher-order effects that would be too complex for paper and pencil analysis. SPICE is nowadays available in a variety of versions undergoing constant revision, so rather than committing to a particular version, I have decided to keep the examples simple

enough for students to quickly redraw them and run them in the SPICE version of their choice.

As in the previous editions, the presentation is enhanced by carefully thought-out examples and end-of-chapter problems emphasizing intuition, physical insight, and problem-solving methodologies of the type engineers exercise daily on the job.

The desire to address general and lasting principles in a manner that transcends the latest technological trend has motivated the choice of well-established and widely documented devices as vehicles. However, when necessary, students are made aware of more recent alternatives, which they are encouraged to look up online.

THE CONTENTS AT A GLANCE

Although not explicitly indicated, the book consists of three parts. The first part (Chapters 1–4) introduces fundamental concepts and applications based on the op amp as a predominantly ideal device. It is felt that the student needs to develop sufficient confidence with ideal (or near-ideal) op amp situations before tackling and assessing the consequences of practical device limitations. Limitations are the subject of the second part (Chapters 5–8), which covers the topic in more systematic detail than previous editions. Finally, the third part (Chapters 9–13) exploits the maturity and judgment developed by the reader in the first two parts to address a variety of design-oriented applications. Following is a brief chapter-by-chapter description of the material covered.

Chapter 1 reviews basic amplifier concepts, including negative feedback. Much emphasis is placed on the loop gain as a gauge of circuit performance. The loop gain is treated via both two-port analysis and return-ratio analysis, with due attention to similarities as well as differences between the two approaches. The student is introduced to simple PSpice models, which will become more sophisticated as we progress through the book. Those instructors who find the loop-gain treatment overwhelming this early in the book may skip it and return to it at a more suitable time. Coverage rearrangements of this sort are facilitated by the fact that individual sections and chapters have been designed to be as independent as possible from each other; moreover, the end-of-chapter problems are grouped by section.

Chapter 2 deals with I-V, V-I, and I-I converters, along with various instrumentation and transducer amplifiers. The chapter places much emphasis on feedback topologies and the role of the loop gain T.

Chapter 3 covers first-order filters, audio filters, and popular second-order filters such as the *KRC,* multiple-feedback, state-variable, and biquad topologies. The chapter emphasizes complex-plane systems concepts and concludes with filter sensitivities.

The reader who wants to go deeper into the subject of filters will find Chapter 4 useful. This chapter covers higher-order filter synthesis using both the cascade and the direct approaches. Moreover, these approaches are presented for both the case of active *RC* filters and the case of switched-capacitor (SC) filters.

Chapter 5 addresses input-referrable op amp errors such as V_{OS}, I_B, I_{OS}, CMRR, PSRR, and drift, along with operating limits. The student is introduced to data-sheet interpretation, PSpice macromodels, and also to different technologies and topologies.

Chapter 6 addresses dynamic limitations in both the frequency and time domains, and investigates their effect on the resistive circuits and the filters that were studied in the first part using mainly ideal op amp models. Voltage feedback and current feedback are compared in detail, and PSpice is used extensively to visualize both the frequency and transient responses of representative circuit examples. Having mastered the material of the first four chapters using ideal or nearly ideal op amps, the student is now in a better position to appreciate and evaluate the consequences of practical device limitations.

The subject of ac noise, covered in Chapter 7, follows naturally since it combines the principles learned in both Chapters 5 and 6. Noise calculations and estimation represent another area in which PSpice proves a most useful tool.

The second part concludes with the subject of stability in Chapter 8. The enhanced coverage of negative feedback has required an extensive revision of frequency compensation, both internal and external to the op amp. The fourth edition makes generous use of the voltage/current injection techniques pioneered by R. D. Middlebrook for loop-gain measurements. Again, PSpice is used profusely to visualize the effect of the different frequency-compensation techniques presented.

The third part begins with nonlinear applications, which are discussed in Chapter 9. Here, nonlinear behavior stems from either the lack of feedback (voltage comparators), or the presence of feedback, but of the positive type (Schmitt triggers), or the presence of negative feedback, but using nonlinear elements such as diodes and switches (precision rectifiers, peak detectors, track-and-hold amplifiers).

Chapter 10 covers signal generators, including Wien-bridge and quadrature oscillators, multivibrators, timers, function generators, and V-F and F-V converters.

Chapter 11 addresses regulation. It starts with voltage references, proceeds to linear voltage regulators, and concludes with a much-expanded coverage of switching regulators. Great attention is devoted to current control and slope compensation, along with stability issues such as error-amplifier design and the effect of the right-half plane zero in boost converters.

Chapter 12 deals with data conversion. Data-converter specifications are treated in systematic fashion, and various applications with multiplying DACs are presented. The chapter concludes with oversampling-conversion principles and sigma-delta converters. Much has been written about this subject, so this chapter of necessity exposes the student only to the fundamentals.

Chapter 13 concludes the book with a variety of nonlinear circuits, such as log/antilog amplifiers, analog multipliers, and operational transconductance amplifiers with a brief exposure to g_m-C filters. The chapter culminates with an introduction to phase-locked loops, a subject that combines important materials addressed at various points in the preceding chapters.

WEBSITE

The book is accompanied by a Website (http://www.mhhe.com/franco) containing information about the book and a collection of useful resources for the instructor. Among the Instructor Resources are a Solutions Manual, a set of PowerPoint Lecture Slides, and a link to the Errata.

ACKNOWLEDGMENTS

Some of the changes in the fourth edition were made in response to feedback received from a number of readers in both industry and academia, and I am grateful to all who took the time to e-mail me. In addition, the following reviewers provided detailed commentaries on the previous edition as well as valuable suggestions for the current revision. All suggestions have been examined in detail, and if only a portion of them has been honored, it was not out of callousness, but because of production constraints or personal philosophy. To all reviewers, my sincere thanks: Aydin Karsilayan, Texas A&M University; Paul T. Kolen, San Diego State University; Jih-Sheng (Jason) Lai, Virginia Tech; Andrew Rusek, Oakland University; Ashok Srivastava, Louisiana State University; S. Yuvarajan, North Dakota State University.

I remain grateful to the reviewers of the previous editions: Stanley G. Burns, Iowa State University; Michael M. Cirovic, California Polytechnic State University-San Luis Obispo; J. Alvin Connelly, Georgia Institute of Technology; William J. Eccles, Rose-Hulman Institute of Technology; Amir Farhat, Northeastern University; Ward J. Helms, University of Washington; Frank H. Hielscher, Lehigh University; Richard C. Jaeger, Auburn University; Franco Maddaleno, Politecnico di Torino, Italy; Dragan Maksimovic, University of Colorado-Boulder; Philip C. Munro, Youngstown State University; Thomas G.Owen, University of North Carolina-Charlotte; Dr. Guillermo Rico, New Mexico State University; Mahmoud F. Wagdy, California State University-Long Beach; Arthur B. Williams, Coherent Communications Systems Corporation; and Subbaraya Yuvarajan, North Dakota State University. Finally, I wish to express my gratitude to Diana May, my wife, for her encouragement and steadfast support.

Sergio Franco
San Francisco, California, 2014

CONTENTS

1

OPERATIONAL AMPLIFIER FUNDAMENTALS

运算放大器（operational amplifier），或者简称 op amp，是在 1947 年由 John R. Ragazzini 命名的，用以代表一种特殊类型的放大器。通过对其外部元件的适当选取，可以构成各种运算，如放大、加、减、微分和积分。运算放大器的首次应用是在模拟计算机中。实现数学运算的能力是由于运算放大器具有高增益并结合负反馈的结果。

早期的运算放大器采用真空二极管实现，因而体积大、耗电并且价格昂贵。第一次显著小型化是由于双极型晶体管（BJT）的出现，促使利用分立 BJT 实现新一代运算放大器。然而，真正的突破出现在集成电路（IC）运算放大器的开发，它的元件是以单片的形式制造在只有针头大小的硅芯片上。第一个这样的器件是在 20 世纪 60 年代初由神童半导体公司（Fairchild Semiconductor

Corporation）的 Robert J. Widlar 研制出的。在 1968 年仙童半导体公司推出了运算放大器从而成为行业标准，这就是流行的 μA741。从那以后，运算放大器的各种系列和制造商急剧涌现。不管怎样，μA741 年无疑是最广泛记载的运算放大器。它的应用普及经久不衰，且当前文献仍有许多拿它作参考，所以无论是从历史观点还是教学角度，它都值得我们学习。

事实上，运算放大器已经持续不断地渗透到模拟和模拟 – 数字混合电子学的每各个方面[1]。如此广泛地应用得益于价格的急剧下降。今天，批量采购一块运算放大器的价格可以与大多传统元件（如微调电容器、质量好的电容器和精密电阻器）的价格相比拟。事实上，一般把运算放大器看作另一种元件，这对当今我们了解和设计模拟电路具深远影响。

在第 5 章的附录中，图 5A.2 为 μA741 运算放大器的内部电路图。这张电路图或许使你感到畏惧，特别是如果你对 BJT 理解还不够深。然而，设计大量的运算放大器应用电路而无须详细了解运算放大器的内部工作机理是可能的。确实如此，无论运算放大器内部多么复杂，它的输入、输出关系却很简单，可以用黑框表示。你将会看到，这种简化的框图对于大多数情况已经够用了。如果不能满足，可以借助于技术数据，并由此预测电路性能，这同样不需要详细考虑内部工作。

为了提升产品，运算放大器制造商一直将应用部门与产品应用效果维系在一起，并且在商业期刊上利用应用笔记和文章将它们公布出来。当今网站上可以获得许多信息，它们鼓励你利用空闲时间，熟悉起模拟产品数据表和应用笔记。你甚至可以签约参加在线研讨会或网络会议。

运算放大器原理的这种学习被实际试验所证实。你可以在实验室里的一块面包板上组装、调试你的电路，也可以采用现在的各种 CAD/CAE 软件包（如 SPICE）用一台个人电脑对它们进行仿真。最好的效果是两者都做。

本章重点

在简要复习基本运放器的概念后，这一章介绍运算放大器，以及适用于研究各种基本运算放大器电路的分析技术，运放电路包括反相 / 非反相放大器、缓冲放大器、加法 / 差分放大器、微分 / 积分器和负阻转换器。

这些电路的工作核心是负反馈技术。本章介绍二端口网络法和返回比分析法。尤其要指出环路增益是负反馈电路最为重要的特性。（在二端口网络法，环路增益和反馈系数分别表示为 L 与 b，在返回比分析法则分别为 T 与 β）。负反馈的优点被大量实例与 SPICE 仿真证实。

本章最后用某些实际应用问题，如运算放大器的供电问题、内部功耗以及输出饱和（更多细节见第 5 章和第 6 章）。本章大量使用 SPICE 作为手算验证工具，同时也用作教学工具来表达更直接的概念和原则。

1.1 AMPLIFIER FUNDAMENTALS

Before embarking on the study of the operational amplifier, it is worth reviewing the fundamental concepts of amplification and loading. Recall that an amplifier is a two-port device that accepts an externally applied signal, called *input,* and generates a signal called *output* such that *output* = *gain* × *input,* where *gain* is a suitable proportionality constant. A device conforming to this definition is called a *linear amplifier* to distinguish it from devices with nonlinear input-output relationships, such as quadratic and log/antilog amplifiers. Unless stated to the contrary, the term *amplifier* will here signify *linear amplifier.*

An amplifier receives its input from a *source* upstream and delivers its output to a *load* downstream. Depending on the nature of the input and output signals, we have different amplifier types. The most common is the *voltage amplifier,* whose input v_I and output v_O are voltages. Each port of the amplifier can be modeled with a Thévenin equivalent, consisting of a voltage source and a series resistance. The input port usually plays a purely passive role, so we model it with just a resistance R_i, called the *input resistance* of the amplifier. The output port is modeled with a voltage-controlled voltage source (VCVS) to signify the dependence of v_O on v_I, along with a series resistance R_o called the *output resistance.* The situation is depicted in Fig. 1.1, where A_{oc} is called the *voltage gain factor* and is expressed in volts per volt. Note that the input source is also modeled with a Thévenin equivalent consisting of the source v_S and an internal series resistance R_s; the output load, playing a passive role, is modeled with a mere resistance R_L.

We now wish to derive an expression for v_O in terms of v_S. Applying the voltage divider formula at the output port yields

$$v_O = \frac{R_L}{R_o + R_L} A_{oc} v_I \tag{1.1}$$

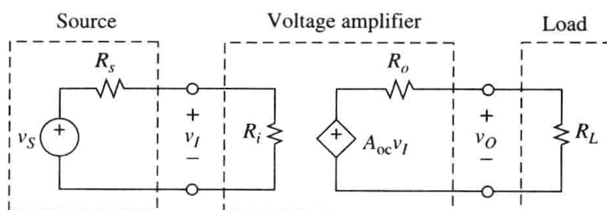


FIGURE 1.1
Voltage amplifier.

We note that in the absence of any load ($R_L = \infty$) we would have $v_O = A_{oc} v_I$. Hence, A_{oc} is called the *unloaded,* or *open-circuit,* voltage gain. Applying the voltage divider formula at the input port yields

$$v_I = \frac{R_i}{R_s + R_i} v_S \tag{1.2}$$

Eliminating v_I and rearranging, we obtain the *source-to-load gain,*

$$\frac{v_O}{v_S} = \frac{R_i}{R_s + R_i} A_{oc} \frac{R_L}{R_o + R_L} \tag{1.3}$$

As the signal progresses from source to load, it undergoes first some attenuation at the input port, then magnification by A_{oc} inside the amplifier, and finally additional attenuation at the output port. These attenuations are referred to as *loading.* It is apparent that because of loading, Eq. (1.3) gives $|v_O/v_S| \le |A_{oc}|$.

EXAMPLE 1.1. (*a*) An amplifier with $R_i = 100$ kΩ, $A_{oc} = 100$ V/V, and $R_o = 1$ Ω is driven by a source with $R_s = 25$ kΩ and drives a load $R_L = 3$ Ω. Calculate the overall gain as well as the amount of input and output loading. (*b*) Repeat, but for a source with $R_s = 50$ kΩ and a load $R_L = 4$ Ω. Compare.

Solution.

(*a*) By Eq. (1.3), the overall gain is $v_O/v_S = [100/(25 + 100)] \times 100 \times 3/(1 + 3) = 0.80 \times 100 \times 0.75 = 60$ V/V, which is less than 100 V/V because of loading. Input loading causes the source voltage to drop to 80% of its unloaded value; output loading introduces an additional drop to 75%.

(*b*) By the same equation, $v_O/v_S = 0.67 \times 100 \times 0.80 = 53.3$ V/V. We now have more loading at the input but less loading at the output. Moreover, the overall gain has changed from 60 V/V to 53.3 V/V.

Loading is generally undesirable because it makes the overall gain dependent on the particular input source and output load, not to mention gain reduction. The origin of loading is obvious: when the amplifier is connected to the input source, R_i draws current and causes R_s to drop some voltage. It is precisely this drop that, once subtracted from v_S, leads to a reduced voltage v_I. Likewise, at the output port the magnitude of v_O is less than the dependent-source voltage $A_{oc} v_I$ because of the voltage drop across R_o.

If loading could be eliminated altogether, we would have $v_O/v_S = A_{oc}$ regardless of the input source and the output load. To achieve this condition, the voltage drops across R_s and R_o must be zero regardless of R_s and R_L. The only way to achieve this is by requiring that our voltage amplifier have $R_i = \infty$ and $R_o = 0$. For obvious reasons such an amplifier is termed *ideal.* Though these conditions cannot be met in practice, an amplifier designer will strive to approximate them as closely as possible by ensuring that $R_i \gg R_s$ and $R_o \ll R_L$ for all input sources and output loads that the amplifier is likely to be connected to.

Another popular amplifier is the *current amplifier.* Since we are now dealing with currents, we model the input source and the amplifier with Norton equivalents, as in Fig. 1.2. The parameter A_{sc} of the current-controlled current source (CCCS)

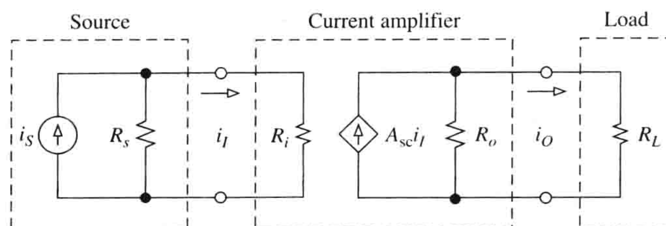


FIGURE 1.2
Current amplifier.

is called the *unloaded,* or *short-circuit, current gain.* Applying the current divider formula twice yields the source-to-load gain,

$$\frac{i_O}{i_S} = \frac{R_s}{R_s + R_i} A_{\text{sc}} \frac{R_o}{R_o + R_L} \tag{1.4}$$

We again witness loading both at the input port, where part of i_S is lost through R_s, making i_I less than i_S, and at the output port, where part of $A_{\text{sc}} i_I$ is lost through R_o. Consequently, we always have $|i_O/i_S| \leq |A_{\text{sc}}|$. To eliminate loading, an *ideal* current amplifier has $R_i = 0$ and $R_o = \infty$, exactly the opposite of the ideal voltage amplifier.

An amplifier whose input is a voltage v_I and whose output is a current i_O is called a *transconductance amplifier* because its gain is in amperes per volt, the dimensions of conductance. The situation at the input port is the same as that of the voltage amplifier of Fig. 1.1; the situation at the output port is similar to that of the current amplifier of Fig. 1.2, except that the dependent source is now a voltage-controlled current source (VCCS) of value $A_g v_I$, with A_g in amperes per volt. To avoid loading, an ideal transconductance amplifier has $R_i = \infty$ and $R_o = \infty$.

Finally, an amplifier whose input is a current i_I and whose output is a voltage v_O is called a *transresistance amplifier,* and its gain is in volts per ampere. The input port appears as in Fig. 1.2, and the output port as in Fig. 1.1, except that we now have a current-controlled voltage source (CCVS) of value $A_r i_I$, with A_r in volts per ampere. Ideally, such an amplifier has $R_i = 0$ and $R_o = 0$, the opposite of the transconductance amplifier.

The four basic amplifier types, along with their ideal input and output resistances, are summarized in Table 1.1.

TABLE 1.1
Basic amplifiers and their ideal terminal resistances

Input	Output	Amplifier type	Gain	R_i	R_o
v_I	v_O	Voltage	V/V	∞	0
i_I	i_O	Current	A/A	0	∞
v_I	i_O	Transconductance	A/V	∞	∞
i_I	v_O	Transresistance	V/A	0	0

1.2 THE OPERATIONAL AMPLIFIER

The operational amplifier is a voltage amplifier with extremely high gain. For example, the popular 741 op amp has a typical gain of 200,000 V/V, also expressed as 200 V/mV. Gain is also expressed in decibels (dB) as $20 \log_{10} 200{,}000 = 106$ dB. The OP77, a more recent type, has a gain of 12 million, or 12 V/μV, or $20 \log_{10}(12 \times 10^6) = 141.6$ dB. In fact, what distinguishes op amps from all other voltage amplifiers is the size of their gain. In the next sections we shall see that the higher the gain the better, or that an op amp would ideally have an infinitely large gain. Why one would want gain to be extremely large, let alone infinite, will become clearer as soon as we start analyzing our first op amp circuits.

Figure 1.3*a* shows the symbol of the op amp and the power-supply connections to make it work. The inputs, identified by the "−" and "+" symbols, are designated *inverting* and *noninverting*. Their voltages with respect to ground are denoted v_N and v_P, and the output voltage as v_O. The arrowhead signifies signal flow from the inputs to the output.

Op amps do not have a 0-V ground terminal. Ground reference is established externally by the power-supply common. The supply voltages are denoted V_{CC} and V_{EE} in the case of bipolar devices, and V_{DD} and V_{SS} in the case of CMOS devices. The typical dual-supply values of ±15 V of the 741 days have been gradually reduced by over a decade, to the point that nowadays supplies of ±1.25 V, or +1.25 V and 0 V, are not uncommon, especially in portable equipment. As we proceed, we shall use a variety of power-supply values, keeping in mind that most principles and applications you are about to learn are not critically dependent on the particular supplies in use. To minimize cluttering in circuit diagrams, it is customary not to show the power-supply connections. However, when we try out an op amp in the lab, we must remember to apply power to make it function.

Figure 1.3*b* shows the equivalent circuit of a properly powered op amp. Though the op amp itself does not have a ground pin, the ground symbol inside its equivalent circuit models the power-supply common of Fig. 1.3*a*. The equivalent circuit includes the *differential* input resistance r_d, the *voltage gain a*, and the *output resistance* r_o. For reasons that will become clear in the next sections, r_d, a, and r_o are referred to

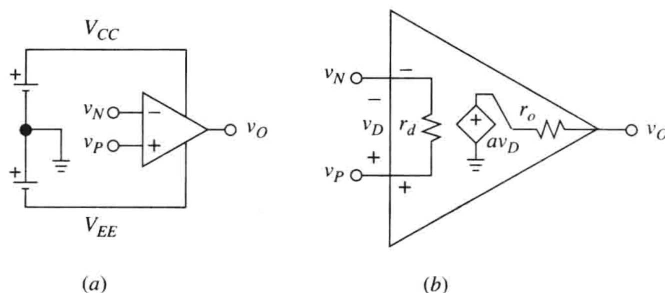


FIGURE 1.3
(*a*) Op amp symbol and power-supply connections. (*b*) Equivalent circuit of a powered op amp. (The 741 op amp has typically $r_d = 2$ MΩ, $a = 200$ V/mV, and $r_o = 75$ Ω.)

as *open-loop* parameters and are symbolized by lowercase letters. The difference

$$v_D = v_P - v_N \tag{1.5}$$

is called the *differential input voltage,* and gain a is also called the *unloaded gain* because in the absence of output loading we have

$$v_O = av_D = a(v_P - v_N) \tag{1.6}$$

Since both input terminals are allowed to attain independent potentials with respect to ground, the input port is said to be of the *double-ended* type. Contrast this with the output port, which is of the *single-ended* type. Equation (1.6) indicates that the op amp responds only to the difference between its input voltages, not to their individual values. Consequently, op amps are also called *difference amplifiers.*

Reversing Eq. (1.6), we obtain

$$v_D = \frac{v_O}{a} \tag{1.7}$$

which allows us to find the voltage v_D causing a given v_O. We again observe that this equation yields only the difference v_D, not the values of v_N and v_P themselves. Because of the large gain a in the denominator, v_D is bound to be very small. For instance, to sustain $v_O = 6$ V, an unloaded 741 op amp needs $v_D = 6/200{,}000 = 30\ \mu$V, quite a small voltage. An unloaded OP77 would need $v_D = 6/(12 \times 10^6) = 0.5\ \mu$V, an even smaller value!

The Ideal Op Amp

We know that to minimize loading, a well-designed voltage amplifier must draw negligible (ideally zero) current from the input source and must present negligible (ideally zero) resistance to the output load. Op amps are no exception, so we define the ideal op amp as an ideal voltage amplifier with infinite open-loop gain:

$$a \to \infty \tag{1.8a}$$

Its ideal terminal conditions are

$$r_d = \infty \tag{1.8b}$$

$$r_o = 0 \tag{1.8c}$$

$$i_P = i_N = 0 \tag{1.8d}$$

where i_P and i_N are the currents drawn by the noninverting and inverting inputs. The ideal op amp model is shown in Fig. 1.4.

We observe that in the limit $a \to \infty$, we obtain $v_D \to v_O/\infty \to 0$! This result is often a source of puzzlement because it makes one wonder how an amplifier with zero input can sustain a nonzero output. Shouldn't the output also be zero by Eq. (1.6)? The answer lies in the fact that as gain a approaches infinity, v_D does indeed approach zero, but in such a way as to maintain the product av_D nonzero and equal to v_O.

Real-life op amps depart somewhat from the ideal, so the model of Fig. 1.4 is only a conceptualization. But during our initiation into the realm of op amp circuits, we shall use this model because it relieves us from worrying about loading effects

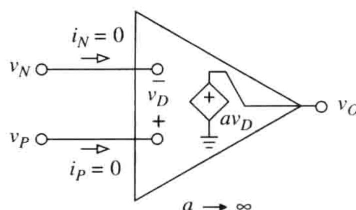


FIGURE 1.4
Ideal op amp model.

so that we can concentrate on the role of the op amp itself. Once we have developed enough understanding and confidence, we shall backtrack and use the more realistic model of Fig. 1.3*b* to assess the validity of our results. We shall find that the results obtained with the ideal and with the real-life models are in much closer agreement than we might have suspected, corroborating the claim that the ideal model, though a conceptualization, is not that academic after all.

SPICE Simulation

Circuit simulation by computer has become a powerful and indispensable tool in both analysis and design. In this book we shall use SPICE, both to verify our calculations and to investigate higher-order effects that would be too complex for paper-and-pencil analysis. The reader is assumed to be conversant with the SPICE basics covered in prerequisite courses. SPICE is available in a wide variety of versions under continuous revision. Though the circuit examples of this book were created using the Student Version of Cadence's PSpice, the reader can easily redraw and rerun them in the version of SPICE in his/her possession.

We begin with the basic model of Fig. 1.5, which reflects 741 data. The circuit uses a voltage-controlled voltage source (VCVS) to model voltage gain, and a resistor pair to model the terminal resistances (by PSpice convention, the "+" input is shown at the top and the "−" input at the bottom, just the opposite of op amp convention).

If a pseudo-ideal model is desired, then r_d is left open, r_o is shorted out, and the source value is increased from 200 kV/V to some huge value, say, 1 GV/V. (However, the reader is cautioned that too large a value may cause convergence problems.)

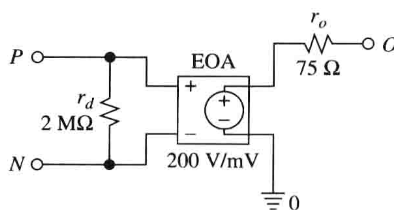


FIGURE 1.5
Basic SPICE model of the 741 op amp.

1.3 BASIC OP AMP CONFIGURATIONS

By connecting external components around an op amp, we obtain what we shall henceforth refer to as an *op amp circuit.* It is crucial that you understand the difference between an op amp circuit and a plain op amp. Think of the latter as a component of the former, just as the external components are. The most basic op amp circuits are the *inverting, noninverting,* and *buffer amplifiers.*

The Noninverting Amplifier

The circuit of Fig. 1.6*a* consists of an op amp and two external resistors. To understand its function, we need to find a relationship between v_O and v_I. To this end we redraw it as in Fig. 1.6*b*, where the op amp has been replaced by its equivalent model and the resistive network has been rearranged to emphasize its role in the circuit. We can find v_O via Eq. (1.6); however, we must first derive expressions for v_P and v_N. By inspection,

$$v_P = v_I \tag{1.9}$$

Using the voltage divider formula yields $v_N = [R_1/(R_1 + R_2)]v_O$, or

$$v_N = \frac{1}{1 + R_2/R_1} v_O \tag{1.10}$$

The voltage v_N represents the fraction of v_O that is being fed back to the inverting input. Consequently, the function of the resistive network is to create *negative feedback* around the op amp. Letting $v_O = a(v_P - v_N)$, we get

$$v_O = a\left(v_I - \frac{1}{1 + R_2/R_1} v_O\right) \tag{1.11}$$

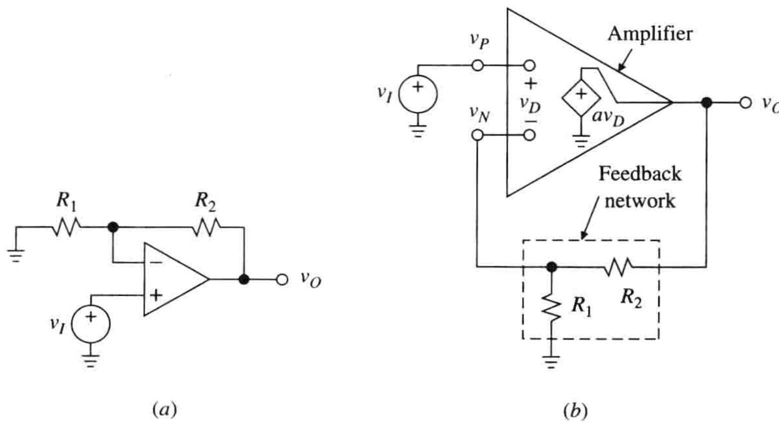


FIGURE 1.6
Noninverting amplifier and circuit model for its analysis.

Collecting terms and solving for the ratio v_O/v_I, which we shall designate as A, yields, after minor rearrangement,

$$A = \frac{v_O}{v_I} = \left(1 + \frac{R_2}{R_1}\right) \frac{1}{1 + (1 + R_2/R_1)/a} \tag{1.12}$$

This result reveals that the circuit of Fig. 1.6*a*, consisting of an op amp plus a resistor pair, is itself an amplifier, and that its gain is A. Since A is positive, the polarity of v_O is the same as that of v_I—hence the name *noninverting amplifier.*

The gain A of the op amp circuit and the gain a of the basic op amp are quite different. This is not surprising, as the two amplifiers, while sharing the same output v_O, have different inputs, namely, v_I for the former and v_D for the latter. To underscore this difference, a is referred to as the *open-loop gain,* and A as the *closed-loop gain,* the latter designation stemming from the fact that the op amp circuit contains a loop. In fact, starting from the inverting input in Fig. 1.6*b*, we can trace a clockwise loop through the op amp and then through the resistive network, which brings us back to the starting point.

EXAMPLE 1.2. In the circuit of Fig. 1.6*a*, let $v_I = 1$ V, $R_1 = 2$ kΩ, and $R_2 = 18$ kΩ. Find v_O if (*a*) $a = 10^2$ V/V, (*b*) $a = 10^4$ V/V, (*c*) $a = 10^6$ V/V. Comment on your findings.

Solution. Equation (1.12) gives $v_O/1 = (1 + 18/2)/(1 + 10/a)$, or $v_O = 10/(1 + 10/a)$. So

(*a*) $v_O = 10/(1 + 10/10^2) = 9.091$ V,
(*b*) $v_O = 9.990$ V,
(*c*) $v_O = 9.9999$ V.

The higher the gain a, the closer v_O is to 10.0 V.

Ideal Noninverting Amplifier Characteristics

Letting $a \to \infty$ in Eq. (1.12) yields a closed-loop gain that we refer to as ideal:

$$A_{\text{ideal}} = \lim_{a \to \infty} A = 1 + \frac{R_2}{R_1} \tag{1.13}$$

In this limit A becomes independent of a, and its value is set exclusively by the *external resistance ratio,* R_2/R_1. We can now appreciate the reason for wanting $a \to \infty$. Indeed, a circuit whose closed-loop gain depends only on a resistance ratio offers tremendous advantages for the designer since it makes it easy to tailor gain to the application at hand. For instance, suppose you need an amplifier with a gain of 2 V/V. Then, by Eq. (1.13), pick $R_2/R_1 = A - 1 = 2 - 1 = 1$; for example, pick $R_1 = R_2 = 100$ kΩ. Do you want $A = 10$ V/V? Then pick $R_2/R_1 = 9$; for example, $R_1 = 20$ kΩ and $R_2 = 180$ kΩ. Do you want an amplifier with variable gain? Then make R_1 or R_2 variable by means of a potentiometer (pot). For example, if R_1 is a fixed 10-kΩ resistor and R_2 is a 100-kΩ pot configured as a variable resistance from 0 Ω to 100 kΩ, then Eq. (1.13) indicates that the gain can be varied over the range 1 V/V $\leq A \leq$ 11 V/V. No wonder it is desirable that $a \to \infty$. It leads to the simpler expression of Eq. (1.13), and it makes op amp circuit design a real snap!

Another advantage of Eq. (1.13) is that gain A can be made as accurate and stable as needed by using resistors of suitable quality. Actually it is not even necessary that the individual resistors be of high quality; it only suffices that their ratio be so. For example, using two resistances that track each other with temperature so as to maintain a constant ratio will make gain A temperature-independent. Contrast this with gain a, which depends on the characteristics of the resistors, diodes, and transistors inside the op amp, and is therefore sensitive to thermal drift, aging, and production variations. This is a prime example of one of the most fascinating aspects of electronics, namely, the ability to implement high-performance circuits using inferior components!

The advantages afforded by Eq. (1.13) do not come for free. The price is the size of gain a needed to make this equation acceptable within a given degree of accuracy (more on this will follow). It is often said that we are in effect throwing away a good deal of open-loop gain for the sake of stabilizing the closed-loop gain. Considering the benefits, the price is well worth paying, especially with IC technology, which, in mass production, makes it possible to achieve high open-loop gains at extremely low cost.

Since the op amp circuit of Fig. 1.6 has proven to be an amplifier itself, besides gain A it must also present input and output resistances, which we shall designate as R_i and R_o and call the *closed-loop input* and *output resistances.* You may have noticed that to keep the distinction between the parameters of the basic op amp and those of the op amp circuit, we are using lowercase letters for the former and uppercase letters for the latter.

Though we shall have more to say about R_i and R_o from the viewpoint of negative feedback in Section 1.6, we presently use the simplified model of Fig. 1.6*b* to state that $R_i = \infty$ because the noninverting input terminal appears as an open circuit, and $R_o = 0$ because the output comes directly from the source av_D. In summary,

$$R_i = \infty \qquad R_o = 0 \tag{1.14}$$

which, according to Table 1.1, represent the ideal terminal charactistics of a voltage amplifier. The equivalent circuit of the ideal noninverting amplifier is shown in Fig. 1.7.

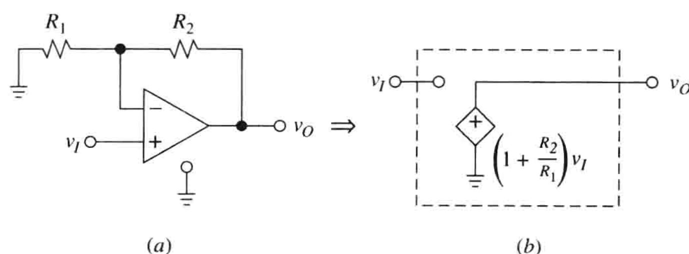


FIGURE 1.7
Noninverting amplifier and its ideal equivalent circuit.

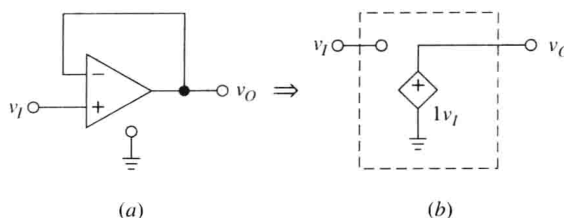


FIGURE 1.8
Voltage follower and its ideal equivalent circuit.

The Voltage Follower

Letting $R_1 = \infty$ and $R_2 = 0$ in the noninverting amplifier turns it into the *unity-gain amplifier,* or *voltage follower,* of Fig. 1.8*a*. Note that the circuit consists of the op amp and a wire to feed the entire output back to the input. The closed-loop parameters are

$$A = 1 \text{ V/V} \qquad R_i = \infty \qquad R_o = 0 \tag{1.15}$$

and the equivalent circuit is shown in Fig. 1.8*b*. As a voltage amplifier, the follower is not much of an achiever since its gain is only unity. Its specialty, however, is to act as a *resistance transformer,* since looking into its input we see an open circuit, but looking into its output we see a short circuit to a source of value $v_O = v_I$.

To appreciate this feature, consider a source v_S whose voltage we wish to apply across a load R_L. If the source were ideal, all we would need would be a plain wire to connect the two. However, if the source has nonzero output resistance R_s, as in Fig. 1.9*a*, then R_s and R_L will form a voltage divider and the magnitude of v_L will be less than that of v_S because of the voltage drop across R_s. Let us now replace the wire by a voltage follower as in Fig. 1.9*b*. Since the follower has $R_i = \infty$, there is no loading at the input, so $v_I = v_S$. Moreover, since the follower has $R_o = 0$, loading is absent also from the output, so $v_L = v_I = v_S$, indicating that R_L now receives the full source voltage with no losses. The role of the follower is thus to act as a *buffer* between source and load.

We also observe that now the source delivers no current and hence no power, while in the circuit of Fig. 1.9*a*, it did. The current and power drawn by R_L are now supplied by the op amp, which in turn takes them from its power supplies,

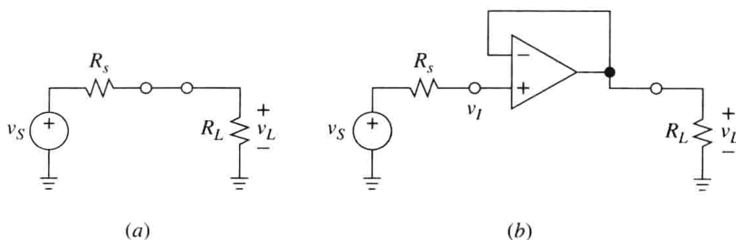


FIGURE 1.9
Source and load connected (*a*) directly, and (*b*) via a voltage follower to eliminate loading.

not explicitly shown in the figure. Thus, besides restoring v_L to the full value of v_S, the follower relieves the source v_S from supplying any power. The need for a buffer arises so often in electronic design that special circuits are available whose performance has been optimized for this function. The BUF03 is a popular example.

The Inverting Amplifier

Together with the noninverting amplifier, the inverting configuration of Fig. 1.10*a* constitutes a cornerstone of op amp applications. The inverting amplifier was invented before the noninverting amplifier because in their early days op amps had only one input, namely, the inverting one. Referring to the equivalent circuit of Fig. 1.10*b*, we have

$$v_P = 0 \tag{1.16}$$

Applying the superposition principle yields $v_N = [R_2/(R_1 + R_2)]v_I + [R_1/(R_1 + R_2)]v_O$, or

$$v_N = \frac{1}{1 + R_1/R_2}v_I + \frac{1}{1 + R_2/R_1}v_O \tag{1.17}$$

Letting $v_O = a(v_P - v_N)$ yields

$$v_O = a\left(-\frac{1}{1 + R_1/R_2}v_I - \frac{1}{1 + R_2/R_1}v_O\right) \tag{1.18}$$

Comparing with Eq. (1.11), we observe that the resistive network still feeds the portion $1/(1 + R_2/R_1)$ of v_O back to the inverting input, thus providing the same amount of negative feedback. Solving for the ratio v_O/v_I and rearranging, we obtain

$$A = \frac{v_O}{v_I} = \left(-\frac{R_2}{R_1}\right)\frac{1}{1 + (1 + R_2/R_1)/a} \tag{1.19}$$

Our circuit is again an *amplifier.* However, the gain A is now *negative,* indicating that the polarity of v_O will be opposite to that of v_I. This is not surprising, because we are now applying v_I to the inverting side of the op amp. Hence, the circuit is called an *inverting amplifier.* If the input is a sine wave, the circuit will introduce a *phase reversal,* or, equivalently, a *180° phase shift.*

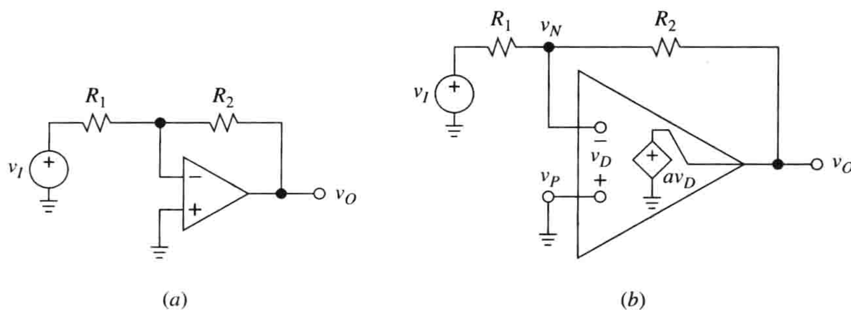


FIGURE 1.10
Inverting amplifier and circuit model for its analysis.

Ideal Inverting Amplifier Characteristics

Letting $a \to \infty$ in Eq. (1.19), we obtain

$$A_{\text{ideal}} = \lim_{a\to\infty} A = -\frac{R_2}{R_1} \tag{1.20}$$

That is, the closed-loop gain again depends only on an external resistance ratio, yielding well-known advantages for the circuit designer. For instance, if we need an amplifier with a gain of -5 V/V, we pick two resistances in a 5:1 ratio, such as $R_1 = 20$ kΩ and $R_2 = 100$ kΩ. If, on the other hand, R_1 is a fixed 20-kΩ resistor and R_2 is a 100-kΩ pot configured as a variable resistance, then the closed-loop gain can be varied anywhere over the range -5 V/V $\le A \le 0$. Note in particular that the magnitude of A can now be controlled all the way down to zero.

We now turn to the task of determining the closed-loop input and output resistances R_i and R_o. Since $v_D = v_O/a$ is vanishingly small because of the large size of a, it follows that v_N is very close to v_P, which is zero. In fact, in the limit $a \to \infty$, v_N would be zero exactly, and would be referred to as *virtual ground* because to an outside observer, things appear as if the inverting input were permanently grounded. We conclude that the effective resistance seen by the input source is just R_1. Moreover, since the output comes directly from the source av_D, we have $R_o = 0$. In summary,

$$R_i = R_1 \qquad R_o = 0 \tag{1.21}$$

The equivalent circuit of the inverting amplifier is shown in Fig. 1.11.

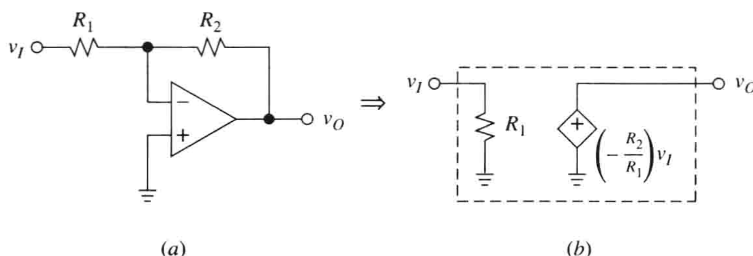


FIGURE 1.11
Inverting amplifier and its ideal equivalent circuit.

EXAMPLE 1.3. (*a*) Using the basic 741 model of Fig. 1.4, direct PSpice to find the closed-loop parameters of an inverting amplifier implemented with $R_1 = 1.0$ kΩ and $R_2 = 100$ kΩ. Compare with the ideal case and comment. (*b*) What happens if you raise a to 1 G V/V? Lower a to 1 kV/V?

Solution.

(*a*) After creating the circuit of Fig. 1.12, we direct PSpice to calculate the small-signal gain `(TF)` from the input source `V(I)` to the output variable `V(O)`. This yields the following output file:

```
V(O)/V(I) = -9.995E+01
INPUT RESISTANCE AT V(I) = 1.001E+03
OUTPUT RESISTANCE AT V(O) = 3.787E-02
```

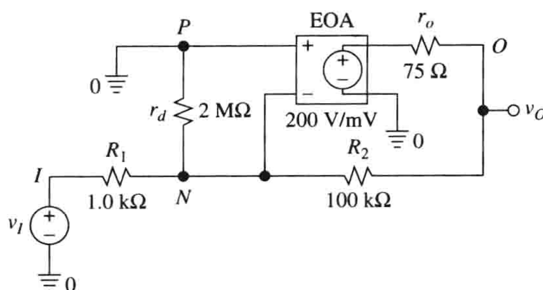


FIGURE 1.12
SPICE circuit for Example 1.3.

It is apparent that the data are quite close to the ideal values $A = -100$ V/V, $R_i = 1.0$ kΩ, and $R_o \to 0$.

(*b*) Rerunning PSpice after increasing the dependent-source gain from 200 kV/V to 1 G V/V we find that A and R_i match the ideal values (within PSpice's resolution) and R_o drops to micro-ohms, which is very close to 0. Lowering the gain to 1 kV/V gives $A = -90.82$ V/V, $R_i = 1.100$ kΩ, and $R_o = 6.883$ Ω, indicating a more pronounced departure from the ideal.

Unlike its noninverting counterpart, the inverting amplifier will load down the input source if the source is nonideal. This is depicted in Fig. 1.13. Since in the limit $a \to \infty$, the op amp keeps $v_N \to 0$ V (virtual ground), we can apply the voltage divider formula and write

$$v_I = \frac{R_1}{R_s + R_1} v_S \tag{1.22}$$

indicating that $|v_I| \leq |v_S|$. Applying Eq. (1.20), $v_L/v_I = -R_2/R_1$. Eliminating v_I, we obtain

$$\frac{v_L}{v_S} = -\frac{R_2}{R_s + R_1}$$ (

Because of loading at the input, the magnitude of the overall gain, $R_2/(R_s + R_1)$, is less than that of the amplifier alone, R_2/R_1. The amount of loading depends on the relative magnitudes of R_s and R_1, and only if $R_s \ll R_1$ can loading be ignored.

We can look at the above circuit also from another viewpoint. Namely, to find the gain v_L/v_S, we can still apply Eq. (1.20), provided, however, that we regard R_s

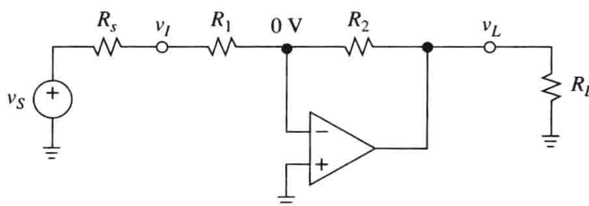


FIGURE 1.13
Input loading by the inverting amplifier.

and R_1 as a *single* resistance of value $R_s + R_1$. Thus, $v_L/v_S = -R_2/(R_s + R_1)$, the same as above.

1.4 IDEAL OP AMP CIRCUIT ANALYSIS

Considering the simplicity of the ideal closed-loop results of the previous section, we wonder whether there is not a simpler technique to derive them, bypassing some of the tedious algebra. Such a technique exists and is based on the fact that when the op amp is operated with negative feedback, in the limit $a \to \infty$ its input voltage $v_D = v_O/a$ approaches zero,

$$\lim_{a\to\infty} v_D = 0 \tag{1.24}$$

or, since $v_N = v_P - v_D = v_P - v_O/a$, v_N approaches v_P,

$$\lim_{a\to\infty} v_N = v_P \tag{1.25}$$

This property, referred to as the *input voltage constraint,* makes the input terminals appear as if they were shorted together, though in fact they are not. We also know that an ideal op amp draws no current at its input terminals, so this apparent short carries no current, a property referred to as the *input current constraint.* In other words, for voltage purposes the input port appears as a short circuit, but for current purposes it appears as an open circuit! Hence the designation *virtual short.* Summarizing, *when operated with negative feedback, an ideal op amp will output whatever voltage and current it takes to drive v_D to zero or, equivalently, to force v_N to track v_P, but without drawing any current at either input terminal.*

Note that it is v_N that tracks v_P, not the other way around. The op amp controls v_N via the external feedback network. Without feedback, the op amp would be unable to influence v_N and the above equations would no longer hold.

To better understand the action of the op amp, consider the simple circuit of Fig. 1.14*a*, where we have, by inspection, $i = 0$, $v_1 = 0$, $v_2 = 6$ V, and $v_3 = 6$ V. If we now connect an op amp as in Fig. 1.14*b*, what will happen? As we know, the op amp will drive v_3 to whatever it takes to make $v_2 = v_1$. To find these voltages,

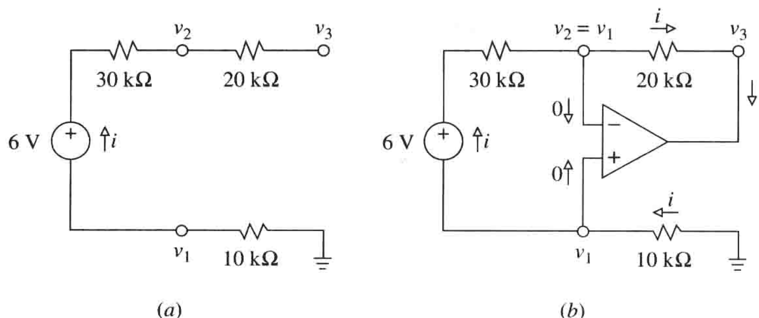


FIGURE 1.14
The effect of an op amp in a circuit.

we equate the current entering the 6-V source to that exiting it; or

$$\frac{0 - v_1}{10} = \frac{(v_1 + 6) - v_2}{30}$$

Letting $v_2 = v_1$ and solving yields $v_1 = -2$ V. The current is

$$i = \frac{0 - v_1}{10} = \frac{2}{10} = 0.2 \text{ mA}$$

and the output voltage is

$$v_3 = v_2 - 20i = -2 - 20 \times 0.2 = -6 \text{ V}$$

Summarizing, as the op amp is inserted in the circuit, it swings v_3 from 6 V to -6 V because this is the voltage that makes $v_2 = v_1$. Consequently, v_1 is changed from 0 V to -2 V, and v_2 from 6 V to -2 V. The op amp also sinks a current of 0.2 mA at its output terminal, but without drawing any current at either input.

The Basic Amplifiers Revisited

It is instructive to derive the noninverting and inverting amplifier gains using the concept of the *virtual short.* In the circuit of Fig. 1.15*a* we exploit this concept to label the inverting-input voltage as v_I. Applying the voltage divider formula, we have $v_I = v_O/(1 + R_2/R_1)$, which is readily turned around to yield the familiar relationship $v_O = (1 + R_2/R_1)v_I$. In words, the noninverting amplifier provides the *inverse* function of the voltage divider: the divider attenuates v_O to yield v_I, whereas the amplifier magnifies v_I by the inverse amount to yield v_O. This action can be visualized via the lever analog depicted above the amplifier in the figure. The lever pivots around a point corresponding to ground. The lever segments correspond to resistances, and the swings correspond to voltages.

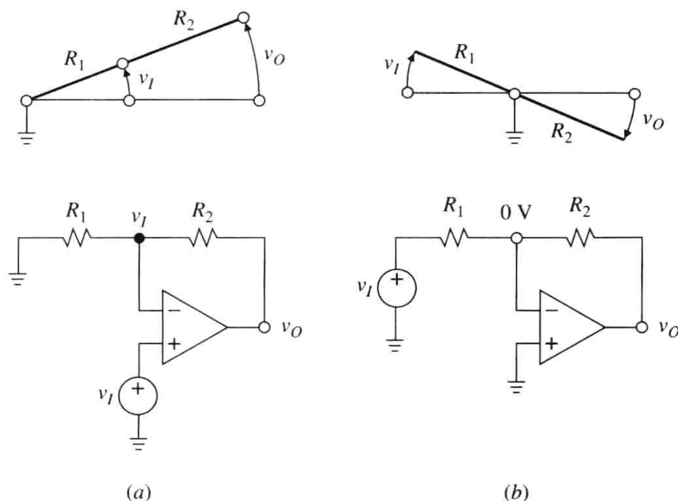


FIGURE 1.15
Mechanical analogies of the noninverting and the inverting amplifiers.

In the circuit of Fig. 1.15*b* we again exploit the virtual-short concept to label the inverting input as a virtual ground, or 0 V. Applying Kirchhoff's current low (KCL), we have $(v_I - 0)/R_1 = (0 - v_O)/R_2$, which is readily solved for v_O to yield the familiar relationship $v_O = (-R_2/R_1)v_I$. This can be visualized via the mechanical analog shown above the amplifier. An upswing (downswing) at the input produces a downswing (upswing) at the output. By contrast, in Fig. 1.15*a* the output swings in the same direction as the input.

So far, we have only studied the basic op amp configurations. It is time to familiarize ourselves with other op amp circuits. These we shall study using the virtual-short concept.

The Summing Amplifier

The summing amplifier has two or more inputs and one output. Though the example of Fig. 1.16 has three inputs, v_1, v_2, and v_3, the following analysis can readily be generalized to an arbitrary number of them. To obtain a relationship between output and inputs, we impose that the total current entering the virtual-ground node equal that exiting it, or

$$i_1 + i_2 + i_3 = i_F$$

For obvious reasons, this node is also referred to as a *summing junction.* Using Ohm's law, $(v_1 - 0)/R_1 + (v_2 - 0)/R_2 + (v_3 - 0)/R_3 = (0 - v_O)/R_F$, or

$$\frac{v_1}{R_1} + \frac{v_2}{R_2} + \frac{v_3}{R_3} = -\frac{v_O}{R_F}$$

We observe that thanks to the virtual ground, the input currents are linearly proportional to the corresponding source voltages. Moreover, the sources are prevented from interacting with each other—a very desirable feature should any of these sources be

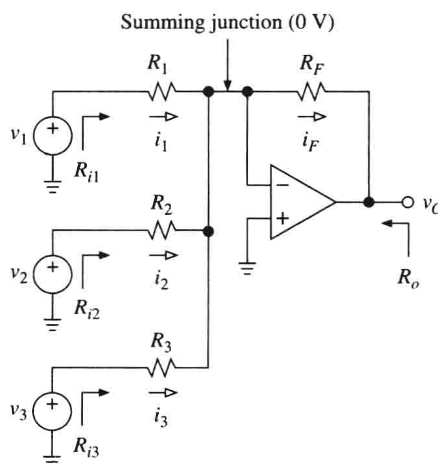


FIGURE 1.16
Summing amplifier.

disconnected from the circuit. Solving for v_O yields

$$v_O = -\left(\frac{R_F}{R_1}v_1 + \frac{R_F}{R_2}v_2 + \frac{R_F}{R_3}v_3\right) \tag{1.26}$$

indicating that the output is a weighted sum of the inputs (hence the name *summing amplifier*), with the weights being established by resistance ratios. A popular application of summing amplifiers is audio mixing.

Since the output comes directly from the dependent source inside the op amp, we have $R_o = 0$. Moreover, because of the virtual ground, the input resistance R_{ik} $(k = 1, 2, 3)$ seen by source v_k equals the corresponding resistance R_k. In summary,

$$R_{ik} = R_k \qquad k = 1, 2, 3$$
$$R_o = 0 \tag{1.27}$$

If the input sources are nonideal, the circuit will load them down, as in the case of the inverting amplifier. Equation (1.26) is still applicable provided we replace R_k by $R_{sk} + R_k$ in the denominators, where R_{sk} is the output resistance of the kth input source.

EXAMPLE 1.4. Using standard 5% resistances, design a circuit such that $v_O = -2(3v_1 + 4v_2 + 2v_3)$.

Solution. By Eq. (1.26) we have $R_F/R_1 = 6$, $R_F/R_2 = 8$, $R_F/R_3 = 4$. One possible standard resistance set satisfying the above conditions is $R_1 = 20$ kΩ, $R_2 = 15$ kΩ, $R_3 = 30$ kΩ, and $R_F = 120$ kΩ.

EXAMPLE 1.5. In the design of function generators and data converters, the need arises to *offset* as well as *amplify* a given voltage v_I to obtain a voltage of the type $v_O = Av_I + V_O$, where V_O is the desired amount of offset. An offsetting amplifier can be implemented with a summing amplifier in which one of the inputs is v_I and the other is either V_{CC} or V_{EE}, the regulated supply voltages used to power the op amp. Using standard 5% resistances, design a circuit such that $v_O = -10v_I + 2.5$ V. Assume ±5-V supplies.

Solution. The circuit is shown in Fig. 1.17. Imposing $v_O = -(R_F/R_1)v_I - (R_F/R_2) \times (-5) = -10v_I + 2.5$, we find that a possible resistance set is $R_1 = 10$ kΩ, $R_2 = 200$ kΩ, and $R_F = 100$ kΩ, as shown.

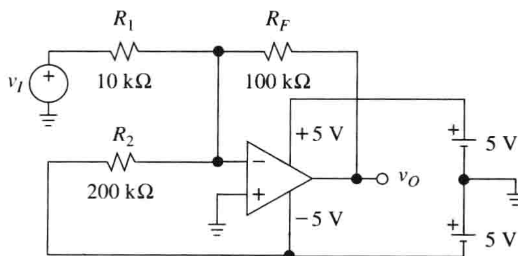


FIGURE 1.17
A dc-offsetting amplifier.

If $R_3 = R_2 = R_1$, then Eq. (1.26) yields

$$v_O = -\frac{R_F}{R_1}(v_1 + v_2 + v_3) \tag{1.28}$$

that is, v_O is proportional to the *true sum* of the inputs. The proportionality constant $-R_F/R_1$ can be varied all the way down to zero by implementing R_F with a variable resistance. If all resistances are equal, the circuit yields the (inverted) sum of its inputs, $v_O = -(v_1 + v_2 + v_3)$.

The Difference Amplifier

As shown in Fig. 1.18, the difference amplifier has one output and two inputs, one of which is applied to the inverting side, the other to the noninverting side. We can find v_O via the superposition principle as $v_O = v_{O1} + v_{O2}$, where v_{O1} is the value of v_O with v_2 set to zero, and v_{O2} that with v_1 set to zero.

Letting $v_2 = 0$ yields $v_P = 0$, making the circuit act as an inverting amplifier with respect to v_1. So $v_{O1} = -(R_2/R_1)v_1$ and $R_{i1} = R_1$, where R_{i1} is the input resistance seen by the source v_1.

Letting $v_1 = 0$ makes the circuit act as a noninverting amplifier with respect to v_P. So $v_{O2} = (1 + R_2/R_1)v_P = (1 + R_2/R_1) \times [R_4/(R_3 + R_4)]v_2$ and $R_{i2} = R_3 + R_4$, where R_{i2} is the input resistance seen by the source v_2. Letting $v_O = v_{O1} + v_{O2}$ and rearranging yields

$$v_O = \frac{R_2}{R_1}\left(\frac{1 + R_1/R_2}{1 + R_3/R_4}v_2 - v_1\right) \tag{1.29}$$

Moreover,

$$R_{i1} = R_1 \qquad R_{i2} = R_3 + R_4 \qquad R_o = 0 \tag{1.30}$$

The output is again a linear combination of the inputs, but with coefficients of opposite polarity because one input is applied to the inverting side and the other to the noninverting side of the op amp. Moreover, the resistances seen by the input sources are finite and, in general, different from each other. If these sources are nonideal, the circuit will load them down, generally by different amounts. Let the

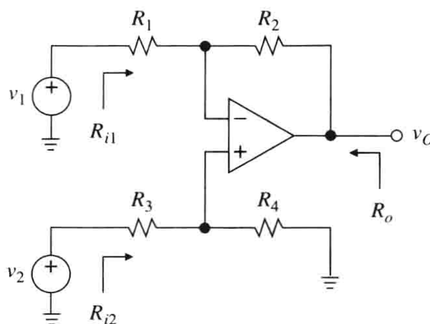


FIGURE 1.18
Difference amplifier.

sources have output resistances R_{s1} and R_{s2}. Then Eq. (1.29) is still applicable provided we replace R_1 by $R_{s1} + R_1$ and R_3 by $R_{s2} + R_3$.

EXAMPLE 1.6. Design a circuit such that $v_O = v_2 - 3v_1$ and $R_{i1} = R_{i2} = 100\ \text{k}\Omega$.

Solution. By Eq. (1.30) we must have $R_1 = R_{i1} = 100\ \text{k}\Omega$. By Eq. (1.29) we must have $R_2/R_1 = 3$, so $R_2 = 300\ \text{k}\Omega$. By Eq. (1.30) $R_3 + R_4 = R_{i2} = 100\ \text{k}\Omega$. By Eq. (1.29), $3[(1 + 1/3)/(1 + R_3/R_4)] = 1$. Solving the last two equations for their two unknowns yields $R_3 = 75\ \text{k}\Omega$ and $R_4 = 25\ \text{k}\Omega$.

An interesting case arises when the resistance pairs in Fig. 1.18 are in equal ratios:

$$\frac{R_3}{R_4} = \frac{R_1}{R_2} \tag{1.31}$$

When this condition is met, the resistances are said to form a *balanced bridge,* and Eq. (1.29) simplifies to

$$v_O = \frac{R_2}{R_1}(v_2 - v_1) \tag{1.32}$$

The output is now proportional to the *true difference* of the inputs—hence the name of the circuit. A popular application of the true difference amplifier is as a building block of instrumentation amplifiers, to be studied in the next chapter.

The Differentiator

To find the input-output relationship for the circuit of Fig. 1.19, we start out by imposing $i_C = i_R$. Using the capacitance law and Ohm's law, this becomes $Cd(v_I - 0)/dt = (0 - v_O)/R$, or

$$v_O(t) = -RC\frac{dv_I(t)}{dt} \tag{1.33}$$

The circuit yields an output that is proportional to the *time derivative* of the input—hence the name. The proportionality constant is set by R and C, and its units are seconds (s).

If you try out the differentiator circuit in the lab, you will find that it tends to oscillate. Its stability problems stem from the open-loop gain rolloff with frequency, an issue that will be addressed in Chapter 8. Suffice it to say here that the circuit is usually stabilized by placing a suitable resistance R_s in series with C. After this

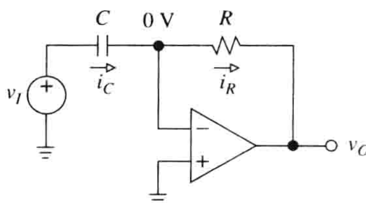


FIGURE 1.19
The op amp differentiator.

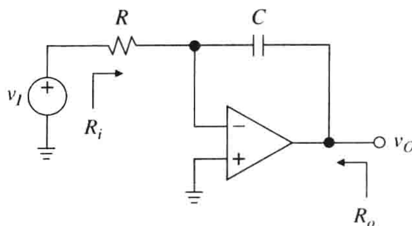


FIGURE 1.20
The op amp integrator.

modification the circuit will still provide the differentiation function, but only over a limited frequency range.

The Integrator

The analysis of the circuit of Fig. 1.20 mirrors that of Fig. 1.19. Imposing $i_R = i_C$, we now get $(v_I - 0)/R = C\,d(0 - v_O)/dt$, or $dv_O(t) = (-1/RC)v_I(t)\,dt$. Changing t to the dummy integration variable ξ and then integrating both sides from 0 to t yields

$$v_O(t) = -\frac{1}{RC}\int_0^t v_I(\xi)\,d\xi + v_O(0) \tag{1.34}$$

where $v_O(0)$ is the value of the output at $t = 0$. This value depends on the charge initially stored in the capacitor. Equation (1.34) indicates that the output is proportional to the *time integral* of the input—hence the name. The proportionality constant is set by R and C, but its units are now s^{-1}. Mirroring the analysis of the inverting amplifier, you can readily verify that

$$R_i = R \qquad R_o = 0 \tag{1.35}$$

Thus, if the driving source has an output resistance R_s, in order to apply Eq. (1.34) we must replace R with $R_s + R$.

The op amp integrator, also called a *precision integrator* because of the high degree of accuracy with which it can implement Eq. (1.34), is a workhorse of electronics. It finds wide application in function generators (triangle and sawtooth wave generators), active filters (state-variable and biquad filters, switched-capacitor filters), analog-to-digital converters (dual-slope converters, quantized-feedback converters), and analog controllers (PID controllers).

If $v_I(t) = 0$, Eq. (1.34) predicts that $v_O(t) = v_O(0) =$ constant. In practice, when the integrator circuit is tried out in the lab, it is found that its output will drift until it saturates at a value close to one of the supply voltages, even with v_I grounded. This is due to the so-called input offset error of the op amp, an issue to be discussed in Chapter 5. Suffice it to say here that a crude method of preventing saturation is to place a suitable resistance R_p in parallel with C. The resulting circuit, called a *lossy integrator,* will still provide the integration function, but only over a limited frequency range. Fortunately, in most applications integrators are placed inside a control loop designed to automatically keep the circuit away from saturation, at least

under proper operating conditions, thus eliminating the need for the aforementioned parallel resistance.

The Negative-Resistance Converter (NIC)

We conclude by demonstrating another important op amp application besides signal processing, namely, *impedance transformation.* To illustrate, consider the plain resistance of Fig. 1.21*a*. To find its value experimentally, we apply a test source v, we measure the current i *out* of the source's positive terminal, and then we let $R_{eq} = v/i$, where R_{eq} is the value of the resistance as seen by the source. Clearly, in this simple case $R_{eq} = R$. Moreover, the test source releases power and the resistance absorbs power.

Suppose we now lift the lower terminal of R off ground and drive it with a noninverting amplifier with the input tied to the other terminal of R, as shown in Fig. 1.21*b*. The current is now $i = [v - (1 + R_2/R_1)v]/R = -R_2 v/(R_1 R)$. Letting $R_{eq} = v/i$ yields

$$R_{eq} = -\frac{R_1}{R_2} R \tag{1.36}$$

indicating that the circuit simulates a *negative resistance.* The meaning of the negative sign is that current is now actually flowing *into* the test source's positive terminal, causing the source to absorb power. Consequently, a negative resistance *releases* power.

If $R_1 = R_2$, then $R_{eq} = -R$. In this case the test voltage v is amplified to $2v$ by the op amp, making R experience a net voltage v, positive at the right. Consequently, $i = -v/R = v/(-R)$.

Negative resistances can be used to neutralize unwanted ordinary resistances, as in the design of current sources, or to control pole location, as in the design of active filters and oscillators.

Looking back at the circuits covered so far, note that by interconnecting suitable components around a high-gain amplifier we can configure it for a variety of *operations:* multiplication by a constant, summation, subtraction, differentiation, integration, and resistance conversion. This explains why it is called *operational!*

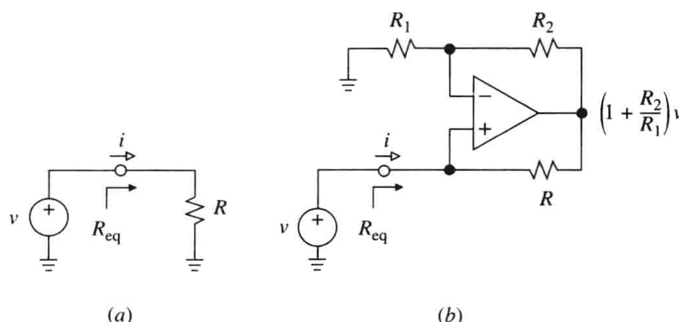


FIGURE 1.21
(*a*) Positive resistance: $R_{eq} = R$. (*b*) Negative-resistance converter: $R_{eq} = -(R_1/R_2)R$.

1.5 NEGATIVE FEEDBACK

Section 1.3 informally introduced the concept of negative feedback. Since the majority of op amp circuits employ this type of feedback, we shall now discuss it in a more systematic fashion.

Figure 1.22 shows the basic structure of a negative-feedback circuit. The arrows indicate signal flow, and the generic symbol x stands for either a voltage or a current signal. Besides the source and load, we identify the following basic blocks:

1. An amplifier called the *error amplifier*, which accepts a signal x_ε called the *error signal*, and yields the *output signal*

$$x_o = a_\varepsilon x_\varepsilon \tag{1.37}$$

   where a_ε is called the *open-loop gain*.
2. A *feedback network*, which samples x_o and produces the *feedback signal*

$$x_f = bx_o \tag{1.38}$$

   where b is the gain of the feedback network and is called the *feedback factor*.
3. A *summing network*, denoted as Σ, which sums the negative of x_f to the *input signal* x_i to yield the difference

$$x_\varepsilon = x_i - x_f \tag{1.39}$$

The designation *negative feedback* stems from the fact that we are in effect feeding a portion b of x_o *back* to the error amplifier's input, where it is subtracted from x_i to yield a *reduced* signal x_ε. Were it added instead, feedback would be positive. For reasons that will become clearer as we move along, negative feedback is also said to be *degenerative*, and positive feedback *regenerative*.

Substituting Eq. (1.38) into Eq. (1.39), and then into Eq. (1.37), we get

$$A = \frac{x_o}{x_i} = \frac{a_\varepsilon}{1 + a_\varepsilon b} \tag{1.40}$$

where A is called the *closed-loop gain* (not to be confused with the open-loop gain $a_\varepsilon = x_o/x_\varepsilon$). Note that for feedback to be negative we must have $a_\varepsilon b > 0$. Consequently, A will be smaller than a_ε by the amount $1 + a_\varepsilon b$, which is called the *amount of feedback*. (Should there be no feedback, we would have $b = 0$ and $A \to a$, a situation referred to as open-loop operation.)

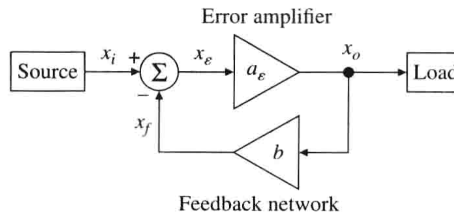


FIGURE 1.22
Block diagram of a negative-feedback system.

As a signal propagates around the loop consisting of the error amplifier, feedback network, and summer, it experiences an overall gain of $a_\varepsilon \times b \times (-1)$, or $-a_\varepsilon b$. Its negative shall be denoted as the *loop gain L*,

$$L = a_\varepsilon b \tag{1.41}$$

This gain allows us to express Eq. (1.40) in the more insightful form $A = (1/b) \times L/(1+L) = (1/b)/(1+1/L)$. Letting $L \to \infty$ yields the ideal situation

$$A_{\text{ideal}} = \lim_{L\to\infty} A = \frac{1}{b} \tag{1.42}$$

that is, A becomes independent of a_ε and is set exclusively by the feedback network, regardless of the error amplifier in use. By proper choice of the topology and components of the feedback network, we can tailor the circuit to a variety of different applications. For instance, specifying $b < 1$ so that $1/b > 1$, will cause x_o to be a *magnified* replica of x_i. Or, implementing the feedback network with reactive elements such as capacitors will yield a *frequency-dependent* circuit with the transfer function $H(s) = 1/b(s)$, where s is the complex frequency. Filters and oscillators are two popular examples.

Henceforth we shall express the closed-loop gain in the insightful form

$$A = A_{\text{ideal}} \frac{1}{1+1/L} \tag{1.43}$$

Rearranging as

$$A = A_{\text{ideal}} \left(1 - \frac{1}{1+L}\right) \tag{1.44}$$

indicates that the *fractional departure* of the actual gain A from the ideal gain A_{ideal} is inversely proportional to the amount of feedback $1 + L$. This departure is more commonly expressed via the *gain error*

$$\text{GE}(\%) = 100\frac{A - A_{\text{ideal}}}{A_{\text{ideal}}} = \frac{1}{1+L} \tag{1.45}$$

EXAMPLE 1.7. (*a*) Find the loop gain needed for GE $\le 0.1\%$. (*b*) Find a_ε to achieve $A = 50$ with the above GE. (*c*) What is the actual value of A? How would you change b to make $A = 50.0$ exactly?

Solution.

(*a*) Use Eq. (1.45) to impose $100/(1 + L) \le 0.1$, or $L \ge 999$ (use $L \ge 10^3$).
(*b*) Use Eq. (1.42) to impose $50 = 1/b$, or $b = 0.02$. Then, $L \ge 10^3 \Rightarrow a_\varepsilon b \ge 10^3 \Rightarrow a_\varepsilon \ge 10^3/0.02 = 5 \times 10^4$.
(*c*) With $L = 10^3$ we have, by Eq. (1.43), $A = 49.95$. To make $A = 50.0$, use Eq. (1.40) to impose $50 = 5 \times 10^4/(1 + 5 \times 10^4 b)$, or $b = 0.01998$.

This example evidences the price for a low gain error, namely, the need to start out with $a_\varepsilon \gg A$ so that, once we close the loop around the error amplifier, we are in effect throwing away a good deal of open-loop gain, namely, the amount of feedback $1 + L$. It is also apparent that for a given a_ε, the smaller the desired gain A, the higher the feedback factor b, and thus the higher the loop gain L and the lower the gain error.

It is instructive to investigate the effect of negative feedback on the signals x_ε and x_f. Writing $x_\varepsilon = x_o/a_\varepsilon = (Ax_i)/a_\varepsilon = (A/a_\varepsilon)x_i$, we get

$$x_\varepsilon = \frac{x_i}{1+L} \tag{1.46}$$

Moreover, writing $x_f = bx_o = b(Ax_i)$ and using Eq. (1.40), we get

$$x_f = \frac{x_i}{1+1/L} \tag{1.47}$$

As $L \to \infty$, the error signal x_ε will approach zero, and the feedback signal x_f will track the input signal x_i. This forms the basis of the already-familiar virtual-short concept.

Gain Desensitivity

We wish to investigate how variations in the open-loop gain a_ε affect the closed-loop gain A. Differentiating Eq. (1.40) with respect to a_ε gives $dA/da_\varepsilon = 1/(1+a_\varepsilon b)^2$. Substituting $1 + a_\varepsilon b = a_\varepsilon/A$ and rearranging,

$$\frac{dA}{A} = \frac{1}{1+L}\frac{da_\varepsilon}{a_\varepsilon}$$

Replacing differentials with finite differences and multiplying both sides by 100, we approximate as

$$100\frac{\Delta A}{A} \cong \frac{1}{1+L}\left(100\frac{\Delta a_\varepsilon}{a_\varepsilon}\right) \tag{1.48}$$

In words, the impact of a given percentage variation in a_ε upon A is reduced by the amount of feedback $1 + L$. So long as L is sufficiently large, even a gross variation in a_ε will cause an insignificant variation in A. We say that negative feedback desensitizes gain, this being the reason why the amount of feedback $1 + L$ is also called the *desensitivity factor*. The stabilization of A is highly desirable because the gain a_ε of a real-life amplifier is ill-defined because of fabrication process variations, thermal drift, and aging.

What about the sensitivity of A to variations in b? Again differentiating Eq. (1.40), but with respect to b, we get

$$100\frac{\Delta A}{A} \cong 100\frac{\Delta b}{b} \tag{1.49}$$

indicating that negative feedback does not stabilize A against variations in b. If we want a stable A, we need to implement the feedback network with components of adequate quality.

EXAMPLE 1.8. A negative-feedback circuit has $a_\varepsilon = 10^5$ and $b = 10^{-3}$. (*a*) Estimate the percentage change in A brought about by a ±10% change in a. (*b*) Repeat if $b = 1$.

Solution.

(*a*) The desensitivity factor is $1 + L = 1 + 10^5 \times 10^{-3} = 101$. Thus, a ±10% change in a_ε will cause a percentage change in A 101 times as small; that is, A changes by $\pm 10/101 \cong \pm 0.1\%$.

(*b*) Now the desensitivity increases to $1 + 10^5 \times 1 \cong 10^5$. The percentage change in A is now $\pm 10/10^5 = 0.0001\%$, or one part per million (1 ppm). We note that for a given a_ε, the lower the value of A, the higher the desensitivity because $1 + L = a_\varepsilon/A$.

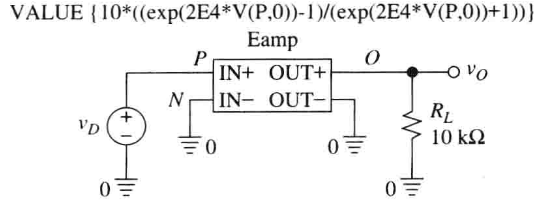


FIGURE 1.23
PSpice amplifier with a nonlinear voltage transfer curve (nonlinear VTC).

Nonlinear Distortion Reduction

A convenient way of visualizing the transfer characteristic of the error amplifier is by means of its *transfer curve*, that is, the plot of the output x_O versus the input x_E. The transfer curve of Eq. (1.37) is a straight line with slope a_ε. However, the transfer curve of a real-life amplifier is generally nonlinear. As an example, let us use the PSpice circuit of Fig. 1.23 to display the function

$$v_O = V_o \tanh \frac{v_E}{V_d} = V_o \frac{\exp(2v_E/V_\varepsilon) - 1}{\exp(2v_E/V_\varepsilon) + 1} \tag{1.50}$$

which mimics qualitatively a real-life *voltage transfer curve* (VTC). Using $V_o = 10$ V and $V_\varepsilon = 100\ \mu$V, and directing PSpice to perform a dc sweep, we obtain the VTC of Fig. 1.24*a*, top. The voltage gain a_ε, visualized as the *slope* of the VTC,

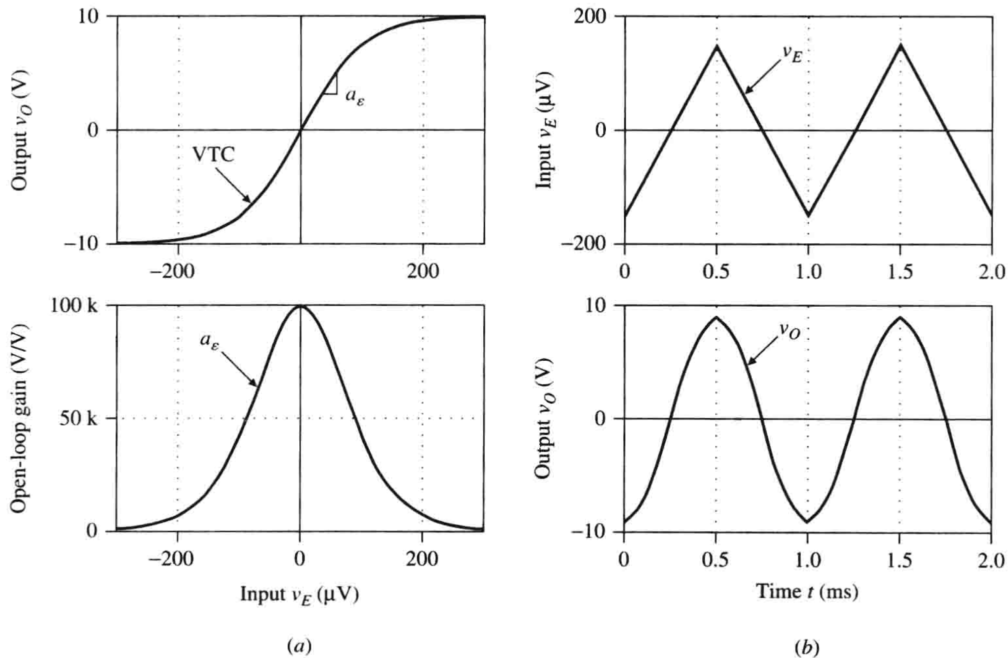


FIGURE 1.24
(*a*) Open-loop characteristics (VTC and gain a_ε) of the amplifier of Fig. 1.23. (*b*) The response v_O to a triangular input v_E.

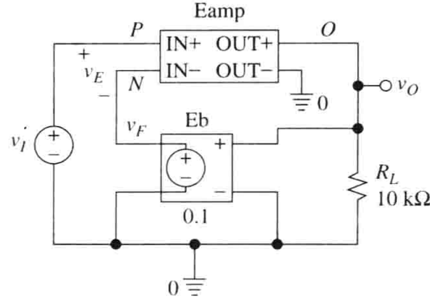


FIGURE 1.25
Applying negative feedback ($b = 0.1$) around the amplifier of Fig. 1.23.

is plotted by PSpice as the function **D(V(O))**. The result, shown at the bottom of Fig. 1.24*a*, indicates that a_ε is maximum at the origin, but decreases as we move away from the origin, and eventually drops to zero at the periphery, where the VTC saturates at ± 10 V.

An obvious consequence of nonlinear gain is *distortion*, as depicted in Fig. 1.24*b* for the case of a triangular input. So long as we keep v_E sufficiently small (say, $\pm 10\ \mu$V or less), v_O will be a relatively undistorted, amplified version of v_E. However, increasing v_E increases the amount of output distortion, as exemplified in the picture. Increasing v_E further would lead to severe output *clipping* and thus even greater distortion.

What are we to make of an amplifier with the above nonlinearity/distortion shortcomings? Enter negative feedback, as depicted in Fig. 1.25, where we use the dependent source **Eb** to sample the output-port voltage v_O, attenuate it to 1/10 to create the feedback signal v_F, and present v_F to the input port of the amplifier, where the latter subtracts it from v_I to create the error signal v_E (note that the operation of input summation is performed inherently by the amplifier itself). The results of the simulation, displayed in Fig. 1.26, reveal a dramatic linearization of the closed-loop VTC: the gain A is fairly constant ($A \cong 1/b = 10$ V/V), and it is so over a much wider output range than the highly nonlinear gain a_ε; moreover, v_O is a fairly undistorted, amplified version of v_I. This, of course, holds over the voltage range for which a_ε is sufficiently large to make the loop gain $L \gg 1$, as per Eqs. (1.42) and (1.43). As we approach the regions of open-loop saturation, where a_ε falls, the linearizing effect of negative feedback no longer applies because of the lack of sufficient loop gain ($L = a_\varepsilon b$) there, so A itself falls.

Given the highly nonlinear VTC of Fig. 1.24*a*, how does the amplifier manage to yield the undistorted waveform v_O of Fig. 1.26*b*? The answer is provided by the error signal v_E, also shown in Fig. 1.26*b*, which illustrates how negative feedback forces the amplifier to *predistort* (via the feedback network) its own input v_E in such a way as to yield the undistorted output v_O. In fact, it was the desire to reduce output distortion that motivated the negative-feedback concept in the first place!

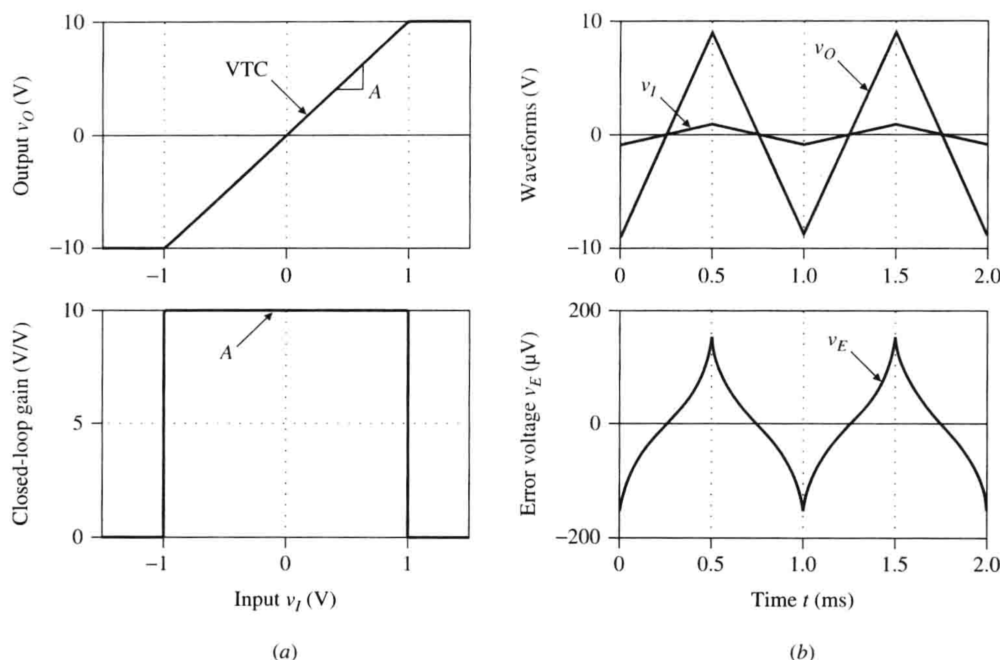


FIGURE 1.26
(a) Closed-loop characteristics (VTC and gain A) of the amplifier of Fig. 1.24. (b) Input v_I, undistorted output v_O, and predistorted error signal v_E.

Effect of Feedback on Disturbances and Noise

Negative feedback provides a means also for reducing circuit sensitivity to certain types of disturbances. Figure 1.27 illustrates three types of disturbances: x_1, entering the circuit at the input, might represent unwanted signals such as input offset errors and input noise, both of which will be covered in later chapters; x_2, entering the circuit at some intermediate point, might represent power-supply hum; x_3, entering the circuit at the output, might represent output load changes.

To accommodate x_2, we break the amplifier into two stages with individual gains $a_{\varepsilon 1}$ and $a_{\varepsilon 2}$, as shown. The overall forward gain is then $a_\varepsilon = a_{\varepsilon 1} a_{\varepsilon 2}$. The output is

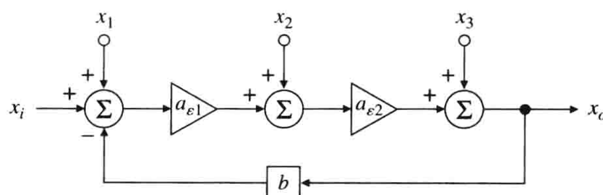


FIGURE 1.27
Investigating the effect of negative feedback on disturbances and noise.

found as $x_o = x_3 + a_{\varepsilon 2}[x_2 + a_{\varepsilon 1}(x_i - bx_o + x_1)]$, or

$$x_o = \frac{a_{\varepsilon 1}a_{\varepsilon 2}}{1 + a_{\varepsilon 1}a_{\varepsilon 2}b}\left(x_i + x_1 + \frac{x_2}{a_{\varepsilon 1}} + \frac{x_3}{a_{\varepsilon 1}a_{\varepsilon 2}}\right)$$

We observe that x_1 undergoes no attenuation relative to x_i. However, x_2 and x_3 are attenuated by the forward gains from the input to the points of entry of the disturbances themselves. This feature is widely exploited in the design of audio amplifiers. The output stage of such an amplifier is a power stage that is usually afflicted by an intolerable amount of hum. Preceding such a stage by a high-gain, low-noise preamplifier stage and then closing a proper feedback loop around the composite amplifier reduces hum by the first-stage gain.

For $a_{\varepsilon 1}a_{\varepsilon 2}b \gg 1$, the above expression simplifies to

$$x_o \cong \frac{1}{b}\left(x_i + x_1 + \frac{x_2}{a_{\varepsilon 1}} + \frac{x_3}{a_{\varepsilon 1}a_{\varepsilon 2}}\right) \tag{1.51}$$

indicating that $1/b$ represents the gain by which the negative-feedback system amplifies the input noise component x_1. Hence, the reason why $1/b$ is often referred to as the *noise gain*.

1.6 FEEDBACK IN OP AMP CIRCUITS

We now wish to relate the concepts of the previous section to circuits based on op amps. Even though the op amp is strictly speaking a *voltage amplifier*, with negative feedback we can operate it as any of the four amplifier types discussed in Section 1.1, giving further credence to its versatility. Accordingly, we have four basic feedback topologies, which pretty much form the basis of all op amp circuits. Our strategy here is to express the signals of each topology in the form

$$x_O = a_\varepsilon(x_I - bx_O) \tag{1.52}$$

so as to identify the gain a_ε and the feedback factor b. In so doing we shall see that the gain a_ε of the feedback circuit may not necessarily coincide with the gain a of the op amp. Once we know a_ε and b, we easily find the loop gain L and the closed-loop gain A via Eqs. (1.41) through (1.43). To help the reader develop an intuitive feel, we start out with the basic op amp model of Fig. 1.4, having $r_d \to \infty$, $r_o \to 0$ and a large gain a.

The Series-Shunt and Shunt-Shunt Topologies

The circuit of Fig. 1.28*a* uses the voltage divider R_1-R_2 to sample v_O and produce the feedback voltage v_F. The op amp then sums $-v_F$ to v_I to produce the error voltage v_D. Since the input-port voltages are summed in series, and the output-port voltage is sampled in parallel (or shunt), this topology is said to be of the *series-shunt* type. By inspection,

$$v_O = av_D = a\,(v_I - v_F) = a\left(v_I - \frac{R_1}{R_1 + R_2}v_F\right)$$

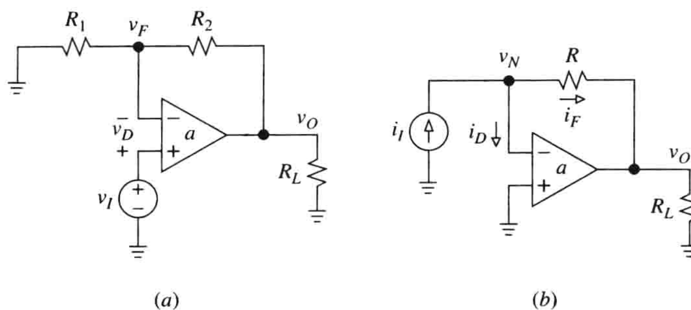


FIGURE 1.28
The (*a*) series-shunt and (*b*) shunt-shunt topologies.

This relationship is of the type of Eq. (1.52), provided

$$a_\varepsilon = a \qquad b = \frac{R_1}{R_1 + R_2} = \frac{1}{1 + R_2/R_1} \tag{1.53}$$

Consequently, the loop gain is

$$L = a_\varepsilon b = \frac{a}{1 + R_2/R_1} \tag{1.54a}$$

Moreover, by Eqs. (1.42) and (1.43), the closed-loop *voltage gain* takes on the form

$$A_v = \frac{v_O}{v_I} = \frac{1}{b}\frac{1}{1 + 1/L} = \left(1 + \frac{R_2}{R_1}\right)\frac{1}{1 + 1/L} \tag{1.54b}$$

These are familiar noninverting-amplifier expressions, but derived from a negative-feedback perspective. As we know, when used as a unity-gain voltage follower, the circuit has $b = 1$, so $L = a$.

The circuit of Fig. 1.28*b* uses the resistance R to sample v_O and to establish the feedback current i_F, whose negative is then summed to i_I to produce the error current i_D. Since the input-port currents are summed in parallel (or shunt), and the output-port voltage is sampled also in parallel, this topology is said to be of the *shunt-shunt* type. Using the superposition principle,

$$v_O = -av_N = -a\,(Ri_I + v_O) = -aR\left(i_I - \frac{-1}{R}v_O\right)$$

This relationship is of the type of Eq. (1.52), provided

$$a_\varepsilon = -aR \qquad b = -\frac{1}{R} \tag{1.55}$$

Note that the open-loop gain $a_\varepsilon (\neq a)$ is now in V/A, and the feedback factor b is in A/V. Both parameters are negative, and their dimensions are reciprocal of each other, so the loop gain

$$L = a_\varepsilon b = a \tag{1.56a}$$

comes out *positive* and *dimensionless*, as it should be (you can exploit this fact to perform a dimensional check on your future derivations). The closed-loop

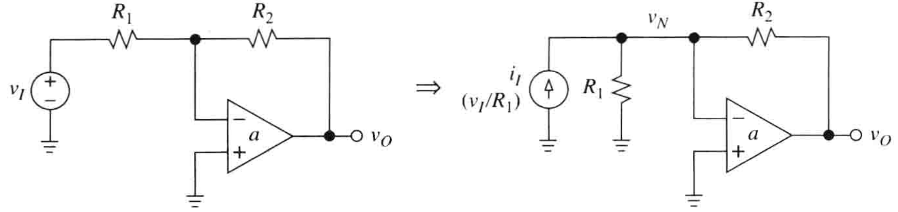


FIGURE 1.29
Performing a source transformation to put the inverting amplifier in shunt-shunt form.

transresistance gain is, by Eqs. (1.42) and (1.43)

$$A_r = \frac{v_O}{i_I} = \frac{1}{b}\frac{1}{1+1/L} = -R\frac{1}{1+1/L} \qquad (1.56b)$$

This gain too is in V/A.

The shunt-shunt topology is at the basis of the popular *inverting voltage amplifier*. This becomes clearer if we perform the input source transformation of Fig. 1.29 by letting

$$i_I = \frac{v_I}{R_1} \qquad (1.57)$$

Then, using the superposition principle, we write

$$v_O = -av_N = -a\left[(R_1//R_2)\,i_I + \frac{R_1}{R_1+R_2}v_O\right] = -a\,(R_1//R_2)\left(i_I - \frac{1}{-R_2}v_O\right)$$

This relationship is of the type of Eq. (1.52), provided

$$a_\varepsilon = -a\,(R_1//R_2) \qquad b = -\frac{1}{R_2} \qquad (1.58)$$

(note that we have again $a_\varepsilon \neq a$). Consequently, the expression for L changes from that of Eq. (1.56*a*) to

$$L = a_\varepsilon b = \left(-a\frac{R_1R_2}{R_1+R_2}\right)\times\left(-\frac{1}{R_2}\right) = \frac{a}{1+R_2/R_1} \qquad (1.59a)$$

Moreover, the closed-loop *transresistance gain* is, by Eqs. (1.42) and (1.43),

$$A_r = \frac{v_O}{i_I} = \frac{1}{b}\frac{1}{1+1/L} = -R_2\frac{1}{1+1/L}$$

whereas the closed-loop *voltage gain* is

$$A_v = \frac{v_O}{v_I} = \frac{v_O}{i_I}\times\frac{i_I}{v_I} = \frac{v_O}{i_I}\times\frac{1}{R_1} = \left(-\frac{R_2}{R_1}\right)\frac{1}{1+1/L} \qquad (1.59b)$$

It is interesting that while the ideal closed-loop *voltage gains* of the inverting and the noninverting configurations are so different, the loop gain L is the *same*, as per Eqs. (1.54*a*) and (1.59*a*). This is so because L is an inherent circuit characteristic established solely by the amplifier and its feedback network.

The Series-Series and Shunt-Series Topologies

The circuit of Fig. 1.30*a* uses the resistance R in series with the load R_L to sample the output-port current i_O and generate the feedback voltage v_F, whose negative is then summed to v_I to produce the error voltage v_D. For obvious reasons, this topology is said to be of the *series-series* type. By inspection,

$$v_O = av_D = a(v_I - v_F) = -a\,(v_I - Ri_O)$$

Substituting $v_O = (R + R_L)i_O$ and rearranging, we get

$$i_O = \frac{a}{R + R_L}(v_I - Ri_O)$$

This relationship is of the type of Eq. (1.52), provided

$$a_\varepsilon = \frac{a}{R + R_L} \qquad b = R \tag{1.60}$$

Note that $a_\varepsilon (\neq a)$ is now in A/V and b in V/A, so the loop gain is dimensionless, as it should be,

$$L = a_\varepsilon b = \frac{a}{1 + R_L/R} \tag{1.61a}$$

Finally, the closed-loop *transconductance* gain, also in A/V, is

$$A_g = \frac{i_O}{v_I} = \frac{1}{b}\frac{1}{1 + 1/L} = \frac{1}{R}\frac{1}{1 + 1/L} \tag{1.61b}$$

The circuit of Fig. 1.30*b* is similar to that of Fig. 1.28*b* at the input port, and to that of Fig. 1.30*a* at the output port, so it is said to be of the *shunt-series* type. It is left as an exercise (see Problem 1.53) to prove that for $a \gg 1$ we have

$$a_\varepsilon \cong -a\frac{1 + R_2/R_1}{1 + R_L/R_1} \qquad b \cong -\frac{1}{1 + R_2/R_1} \tag{1.62}$$

Consequently, the loop gain and the closed-loop *current gain* are, respectively,

$$L \cong \frac{a}{1 + R_L/R_1} \qquad A_i = \frac{i_O}{i_I} = \frac{1}{b}\frac{1}{1 + 1/L} \cong -\left(1 + \frac{R_2}{R_1}\right)\frac{1}{1 + 1/L} \tag{1.63}$$

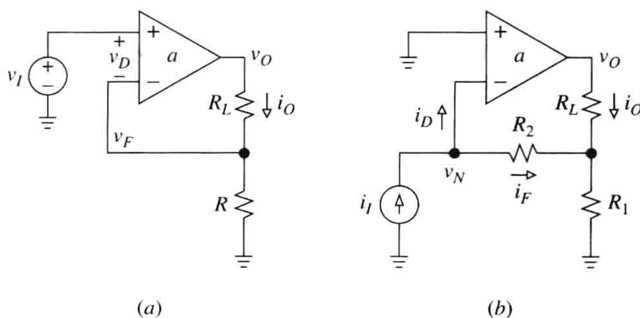


FIGURE 1.30
The (*a*) series-series and (*b*) shunt-series topologies.

(It is interesting that in both circuits, L depends on the particular load R_L in use. Can you justify intuitively why L decreases as R_L is increased?)

EXAMPLE 1.9. In the shunt-series circuit of Fig. 1.30*b* let $R_2 = 2R_1 = 20\ \text{k}\Omega$ and let the op amp have $a = 100$ V/V. (*a*) Calculate a_ε, b, L, and A_i if $R_L = 10\ \text{k}\Omega$. (*b*) Repeat if $R_L = 0$ and comment.

Solution.

(*a*) We have $a_\varepsilon = -100 \times (1 + 20/10)/(1 + 10/10) = -150$ A/A, $b = -1/(1 + 20/10) = -(1/3)$ A/A, $L = 150/3 = 50$, $A_{\text{ideal}} = 1/b = 3.0$ A/A, $A_i = -3/(1 + 1/50) = -2.9412$ A/A.

(*b*) Now $a_\varepsilon = -100 \times (1 + 20/10)/(1 + 0/10) = -300$ A/A, $b = -(1/3)$ A/A, $L = 100$, $A_i = -3/(1 + 1/100) = -2.9703$ A/A. While b remains the same, a_ε increases because of the reduced output swing required of the op amp, causing L to double and thus make A_i closer to ideal.

Closed-Loop Input/Output Resistances

Negative feedback has a dramatic effect not only on gain but also on the terminal resistances (in Chapter 6 we shall see that it affects also the frequency/time responses, and in Chapter 8 the circuit stability).

Turning first to the series-input topology of Fig. 1.31*a*, we observe that owing to feedback, the voltage v_d arising in response to a test voltage v_i is bound to be very small. In fact, Eq. (1.46) gives

$$i_i = \frac{v_d}{r_d} = \frac{v_i}{(1 + L)r_d}$$

so the closed-loop input resistance, obtained as $R_i = v_i/i_i$, is

$$R_i \cong r_d(1 + L) \tag{1.64}$$

(note the use of uppercase letters to distinguish closed-loop parameters from the lowercase-lettered open-loop parameters). In words, negative feedback takes the input resistance r_d, already high in a well-designed op amp, and raises it further by the amount of feedback $1 + L$, which is also large. Clearly, R_i is bound to be much higher than the other resistances in the circuit, so we feel justified assuming $R_i \to \infty$, at least so long as the amount of feedback is sufficiently large.

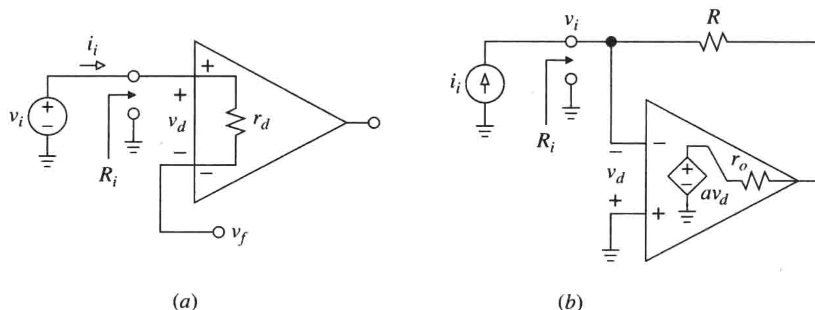


FIGURE 1.31
Finding the closed-loop resistances of the (*a*) series-input and (*b*) shunt-input topologies.

In the shunt-input topology of Fig. 1.31*b* we omit showing r_d because the small voltage v_d across it renders its current insignificant. We feel confident in assuming that the input test current i_i flows entirely into the feedback resistance R, so

$$i_i = \frac{v_i - av_d}{R + r_o} = \frac{v_i - a(-v_i)}{R + r_o}$$

Taking again the ratio $R_i = v_i/i_i$ we get

$$R_i \cong \frac{R + r_o}{1 + a} = \frac{R + r_o}{1 + L} \tag{1.65}$$

where Eq. (1.56*a*) has been used. In a well-designed op amp circuit we usually have $r_o \ll R$, so it is common to approximate $R_i \cong R/(1 + a)$. In words, the feedback resistance R, reflected to the input, gets divided by $1 + a$, which in this case coincides with the amount of feedback. This transformation, known as the *Miller effect*, holds for any feedback impedance, such as capacitive impedances, as we shall see in Chapter 8. With a high gain a we expect $R_i \ll R$. In fact, in the limit $a \to \infty$ we would get $R_i \to 0$, the condition for a *perfect virtual ground*, as we already know. Looking back at the inverting voltage amplifier of Fig. 1.29, we readily find that the resistance seen by the driving source is $v_i/i_i = R_1 + (R_2 + r_o)/(1 + a) \cong R_1$, thus confirming an already familiar result.

We now turn to the closed-loop output resistances, which we find by setting the input source to zero and subjecting the output port to a test signal. With v_i set to 0 V the circuit of Fig. 1.32*a* could have been an inverting or a noninverting amplifier, or it could even be the equivalent circuit of a summing amplifier, a difference amplifier, or a transresistance amplifier. Therefore, the result we are about to derive is going to be quite general. Assume for a moment that i_o flows entirely into r_o, an assumption whose validity we shall verify shortly. We then have, by Kirchhoff's voltage law (KVL) and Ohm's law,

$$v_o \cong av_d + r_o i_o = a\left(-\frac{R_1}{R_1 + R_2}v_o\right) + r_o i_o$$

Collecting and letting $R_o = v_o/i_o$ we get

$$R_o \cong \frac{r_o}{1 + L} \qquad L = \frac{a}{1 + R_2/R_1} \tag{1.66}$$

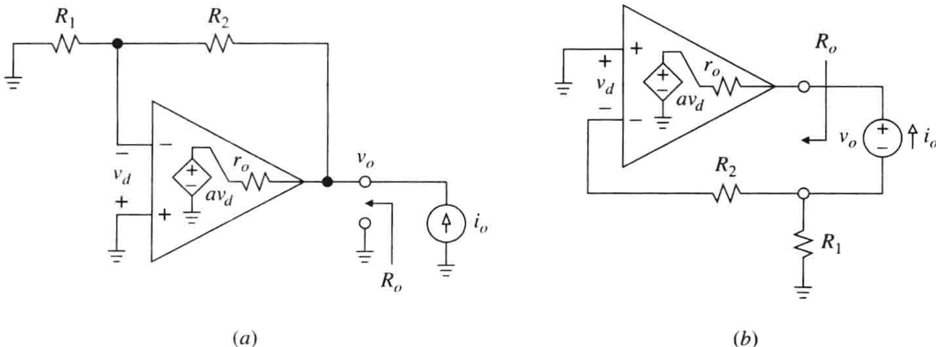


FIGURE 1.32
Finding the closed-loop resistances of the (*a*) shunt-output and (*b*) series-output topologies.

Negative feedback takes the output resistance r_o, already low in a well-designed op amp, and lowers it further by the amount of feedback $1 + L$, which is large. Clearly, R_o is bound to be much lower than the other resistances in the circuit, so we feel justified assuming $R_o \to 0$, at least so long as L is sufficiently large. This also validates our initial assumption that i_o flows almost entirely into the op amp, where it encounters a much smaller resistance than that presented by the feedback network.

Finally, we use the circuit of Fig. 1.32*b* to find the output resistance of a series-output port. With i_i set to 0 V, this circuit could have been the current amplifier of Fig. 1.30*b*, or (with $R_2 = 0$) the transconductance amplifier of Fig. 1.30*a*, or it could be the equivalent circuit of other variants that we shall encounter in the future. Therefore, the forthcoming result is bound to be quite general. Applying the loop method,

$$av_d + r_o i_o - v_o + R_1 i_o = 0 \qquad -v_d + R_1 i_o = 0$$

Eliminating v_d, collecting, and letting $R_o = v_o/i_o$ we get

$$R_o = R_1(1 + a) + r_o \cong R_1(1 + L) \tag{1.67}$$

where we have used Eq. (1.63) with $R_L \to 0$. Clearly, negative feedback increases the output resistance of a series-output port, making the port approach ideal current-source behavior.

EXAMPLE 1.10. (*a*) A 741-type op amp is configured as a noninverting amplifier with $R_1 = 1.0\,\text{k}\Omega$ and $R_2 = 999\,\text{k}\Omega$. Predict R_i, A_v, and R_o, verify with PSpice, and compare with the ideal case. (*b*) Repeat, but with $R_1 = \infty$ and $R_2 = 0$, and comment.

Solution.

(*a*) We have $A_{\text{ideal}} = 1 + 999/1 = 10^3$ V/V and $L = 200{,}000/10^3 = 200$, so $R_i \cong (1 + 200) \times 2 \times 10^6 = 402\ \text{M}\Omega$, $A_v = 10^3/(1 + 1/200) = 995$ V/V, and $R_o = 75/(1 + 200) = 0.373\ \Omega$. These values match those provided by PSpice's small-signal analysis of the circuit of Fig. 1.33. For practical purposes they are fairly close to the ideal values $R_i = \infty$, $A_v = 10^3$ V/V, and $R_o = 0$.

(*b*) The circuit is now a unity-gain voltage follower with $L = a = 2 \times 10^5$, pushing the closed-loop parameters even closer to ideal. Proceeding in similar fashion, we get $R_i = 400\ \text{G}\Omega$, $A_v = 0.999995$ V/V, and $R_o = 0.375\ \text{m}\Omega$, again all in agreement with PSpice.

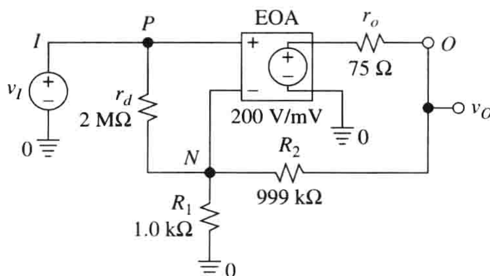


FIGURE 1.33
SPICE circuit for Example 1.10.

It is apparent that the loop gain L plays a vital role in a negative-feedback circuit (we'll have much more to say about this in Chapters 6 and 8). First, L provides a measure of the closed-loop gain's departure from the ideal via Eq. (1.43), repeated here for convenience,

$$A = \frac{x_O}{x_I} = A_{\text{ideal}} \frac{1}{1 + 1/L} \tag{1.68a}$$

where

$$A_{\text{ideal}} = \lim_{T\to\infty} \frac{x_O}{x_I} \tag{1.68b}$$

(In the case of op amp circuits, we derive A_{ideal} using the the input virtual-short concept.) Second, the amount of feedback $1 + L$ represents the amount by which negative feedback raises the resistance of a series-type port as per Eqs. (1.64) and (1.67), or lowers the resistance of a shunt-type port as per Eqs. (1.65) and (1.66). We summarize the input/output resistance transformations as

$$R = r_0(1 + L)^{\pm 1} \tag{1.69}$$

where r_0 is the resistance presented by the port in the limit $L \to 0$ (achieved in the limit $a \to 0$), R is the ensuing closed-loop resistance, and we use $+1$ for series-type ports, -1 for shunt-type ports. The above resistance transformations are highly beneficial in terms of reducing input/output loading. Moreover, they tend to facilitate circuit analysis when the transformed resistances are dramatically larger or smaller than the other resistances in the circuit. The higher L, the closer the closed-loop characteristics to ideal. Put another way, if you had to choose between an op amp with poor r_d and r_o but excellent a, and one with excellent r_d and r_o but poor a, go for the former as the ensuing high L will make up for its poor r_d and r_o characteristics (see Problem 1.60). To stress the importance of L, it is instructive to express a negative-feedback system in terms of L and A_{ideal} rather than in terms of a_ε and b. We do this by letting $b \to 1/A_{\text{ideal}}$ and $a_\varepsilon \to a_\varepsilon b/b = L/b = LA_{\text{ideal}}$, so the block diagram of Fig. 1.22 becomes as shown in Fig. 1.34.

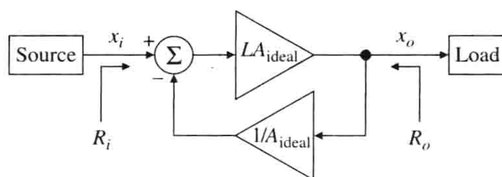


FIGURE 1.34
Negative-feedback system representation in terms of A_{ideal} and the loop gain L.

1.7 THE RETURN RATIO AND BLACKMAN'S FORMULA

The derivations of the previous section postulate (*a*) forward signal transmission exclusively through the error amplifier and (*b*) reverse signal transmission exclusively through the feedback network (the arrowheads of Fig. 1.22 convey explicitly this directionality). Such an amplifier and feedback network are said to be *unilateral*. However, while most op amps do approach unilateral behavior, feedback networks are generally *bilateral*. To develop an intuitive feel for the ensuing implications, we reexamine the familiar inverting and noninverting voltage amplifiers, using the full-blown op amp model of Fig. 1.3*b*.

Summing currents into the nodes labeled v_N and v_O in Fig. 1.35*a* we get

$$\frac{v_I - v_N}{r_d} - \frac{v_N}{R_1} + \frac{v_O - v_N}{R_2} = 0 \qquad \frac{v_N - v_O}{R_2} + \frac{a(v_I - v_N) - v_O}{r_o} = 0$$

where $v_D = v_I - v_N$ has been used. Eliminating v_N, collecting, and solving for the ratio v_O/v_I gives

$$A_{\text{noninv}} = \frac{1 + R_2/R_1}{1 + \frac{1}{a}\left(1 + \frac{R_2}{R_1} + \frac{R_1+r_o}{r_d} + \frac{r_o}{R_1}\right)} + \frac{r_o/r_d}{a + \left(1 + \frac{R_2}{R_1} + \frac{R_1+r_o}{r_d} + \frac{r_o}{R_1}\right)} \tag{1.70}$$

Proceeding in similar fashion with the circuit of Fig. 1.36*a*, we get

$$A_{\text{inv}} = \frac{-R_2/R_1}{1 + \frac{1}{a}\left(1 + \frac{R_2}{R_1} + \frac{R_1+r_o}{r_d} + \frac{r_o}{R_1}\right)} + \frac{r_o/R_1}{a + \left(1 + \frac{R_2}{R_1} + \frac{R_1+r_o}{r_d} + \frac{r_o}{R_1}\right)} \tag{1.71}$$

Both gains can be expressed in the common form[5]

$$A = \frac{A_{\text{ideal}}}{1 + 1/T} + \frac{a_{\text{ft}}}{1 + T} \tag{1.72}$$

where

$$A_{\text{ideal}} = \lim_{a\to\infty} \frac{v_O}{v_I} \tag{1.73}$$

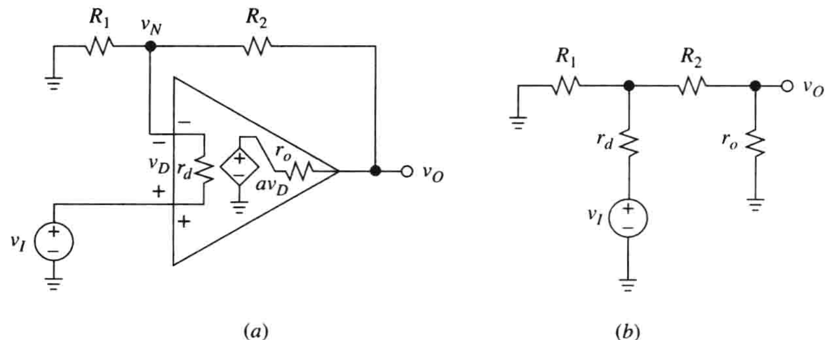


FIGURE 1.35
(*a*) Noninverting amplifier and (*b*) feedthrough signal transmission.

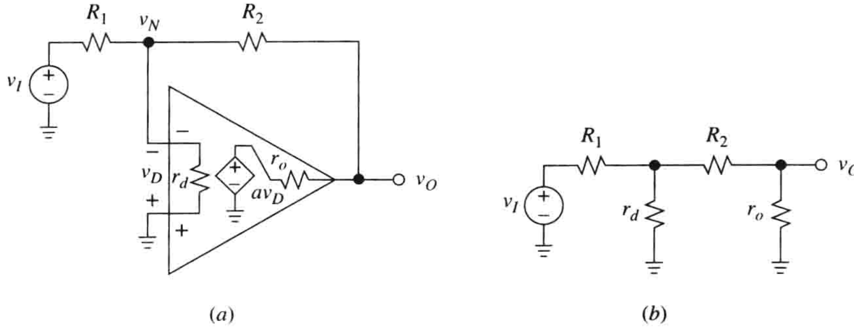


FIGURE 1.36
(*a*) Inverting amplifier and (*b*) feedthrough signal transmission.

is the *ideal closed-loop* gain, calculated via the input virtual-short concept, and

$$a_{\text{ft}} = \lim_{a \to 0} \frac{v_O}{v_I} \tag{1.74}$$

is the *feedthrough gain*, stemming from forward signal transmission around the source av_D, that is, with this source set to zero as in Figs. 1.35*b* and 1.36*b*. You can readily verify that

$$a_{\text{ft(noninv)}} = \frac{r_o/r_d}{1 + R_2/R_1 + (R_1 + r_o)/r_d + r_o/R_1} \tag{1.75a}$$

$$a_{\text{ft(inv)}} = \frac{r_o/R_1}{1 + R_2/R_1 + (R_1 + r_o)/r_d + r_o/R_1} \tag{1.75b}$$

Finally, the quantity

$$T = \frac{a}{1 + R_2/R_1 + (R_1 + r_o)/r_d + r_o/R_1} \tag{1.76}$$

is an important parameter that, for reasons that will be explained below, we shall call the *return-ratio loop gain* (this, to distinguish T from the parameter L of the previous section, which shall henceforth be referred to as the *two-port loop gain*). We visualize Eq. (1.72) as in Fig. 1.37, which you are encouraged to compare with Fig. 1.34 for similarities as well as differences. We wish to make a number of observations:

(*a*) For each circuit a_{ft} is proportional to r_o, so for $r_o \to 0$ both gains vanish as the feedthrough signal gets shunted to the dependent source. Note also that $a_{\text{ft(noninv)}}$ has the same polarity as A_{ideal} whereas $a_{\text{ft(inv)}}$ has the opposite polarity.
(*b*) Since v_I propagates via r_d in the noninverting case and via R_1 in the inverting case, the feedthrough gains are inversely proportional to these resistances. For large r_d we have $a_{\text{ft(noninv)}} << a_{\text{ft(inv)}}$.
(*c*) For $r_d \to \infty$ and $r_o \to 0$, Eq. (1.76) predicts

$$T \to \frac{a}{1 + R_2/R_1} = L \tag{1.77}$$

for both circuits. In a good-quality op amp r_d is large and r_o is small, so for the present circuits T will be somewhat lower than L (though presumably not by much), and a_{ft} will not be that large, indicating that by the time a_{ft} gets divided by $1 + T$, its contribution to A is likely to be negligible for large T.

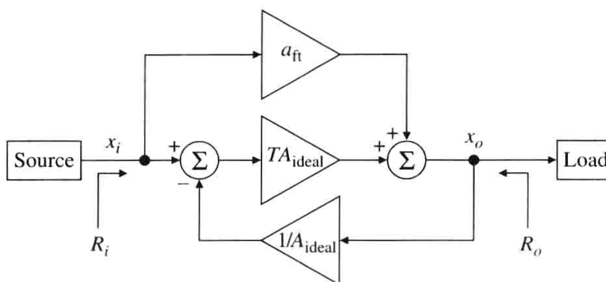


FIGURE 1.37
Negative-feedback system representation in terms of A_{ideal}, T, and a_{ft}.

The Return Ratio of an Op Amp

Though Eq. (1.76) was a by-product of the foregoing analysis, T can be calculated *directly* as follows[2,3]: (*a*) set the input source (or sources in multiple-input circuits such as summing or difference amplifiers) to zero, (*b*) break the circuit right at the dependent source's output, (*c*) inject a test voltage v_T downstream of the source, (*d*) find the voltage av_D returned by the source, and (*e*) obtain T as the negative of the *ratio* of the *returned* voltage to the applied test voltage (hence the designation *return ratio*)

$$T = -\frac{av_D}{v_T} \tag{1.78}$$

(In Chapter 8 we shall discuss how to measure T.) A few examples will better illustrate the procedure.

EXAMPLE 1.11. Repeat Example 1.9, but using return-ratio analysis. Compare and comment.

Solution.

(*a*) The circuit is shown in Fig. 1.38*a*. Setting i_I to zero, breaking the circuit right at the dependent source's output, and injecting a test voltage v_T yields Fig. 1.38*b*. By inspection,

$$av_D = -a\frac{R_1}{R_1 + R_L}v_T \qquad T = -\frac{av_D}{v_T} = \frac{a}{1 + R_L/R_1} = L = \frac{100}{1 + 10/10} = 50$$

Setting the dependent source to zero leads us to Fig. 1.38*c*, where the current divider formula gives

$$a_{\text{ft}} = \frac{i_O}{i_I} = -\frac{R_1}{R_1 + R_L} = -0.5 \text{ A/A}$$

indicating that this circuit has $a_{\text{ft}} \neq 0$ even though $r_o = 0$ has been assumed. As we already know, $A_{\text{ideal}} = -(1 + R_2/R_1) = -3$ A/A, so Eq. (1.72) gives

$$A_i = \frac{-3}{1 + 1/50} + \frac{-0.5}{1 + 50} = -2.9412 - 0.0098 = -2.9510 \text{ A/A}$$

indicating a contribution by a_{ft} of $0.0098/2.9510 = 0.33\%$.

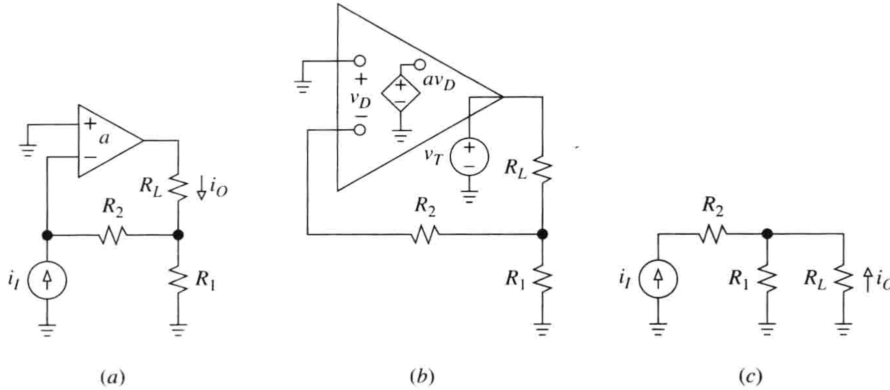


FIGURE 1.38
(*a*) Circuit of Example 1.11. Circuits to find (*b*) T and (*c*) a_{ft}.

(*b*) For $R_L = 0$ we get $T = 100$ and $a_{ft} = -1$ A/A, so

$$A_i = \frac{-3}{1 + 1/100} + \frac{-1}{1 + 100} = -2.9703 - 0.0099 = -2.9802 \text{ A/A}$$

For $R_L = 0$ both T and a_{ft} double, so the percentage contribution due to a_{ft} remains about the same. Compared to the two-port analysis of Example 1.9, return-ratio analysis provides a 0.33% gain correction stemming from the bilateral nature of the feedback network (this, even though $r_d \to \infty$ and $r_o \to 0$ are assumed). Note also that T and L coincide in this particular example.

EXERCISE 1.1. Use return-ratio analysis to verify Eq. (1.76).

EXAMPLE 1.12. (*a*) Perform the return-ratio analysis of the noninverting amplifier of Fig. 1.39*a*, shown in a more general setting that includes also a source resistance R_s and a load resistance R_L. (*b*) Assuming an (mediocre) amplifier with $a = 1000$ V/V, $r_d = 10$ kΩ, and $r_o = 1.0$ kΩ, find the overall source-to-load voltage gain v_L/v_S if $R_1 = 1.0$ kΩ, $R_2 = 9.0$ kΩ, $R_s = 15$ kΩ, and $R_L = 3.0$ kΩ.

Solution.

(*a*) Turning first to Fig. 1.39*b*, we start out at the left and work our way toward the right via repeated application of the voltage divider formula to get

$$av_D = a\frac{-r_d}{R_s + r_d} \times \frac{(R_s + r_d)//R_1}{(R_s + r_d)//R_1 + R_2} \times \frac{[(R_s + r_d)//R_1 + R_2]//R_L}{[(R_s + r_d)//R_1 + R_2]//R_L + r_o} v_T$$

Applying Eq. (1.78), we get, after minor algebra

$$T = a \times \frac{1}{1 + \dfrac{R_s}{r_d}} \times \frac{1}{1 + \dfrac{R_2}{(R_s + r_d)//R_1}} \times \frac{1}{1 + \dfrac{r_o}{[(R_s + r_d)//R_1 + R_2]//R_L}}$$

Next, we start at the right of Fig. 1.39*c* and apply the voltage divider formula twice,

$$v_L = \frac{r_o//R_L}{R_2 + r_o//R_L} \times \frac{R_1//(R_2 + r_o//R_L)}{R_s + r_d + R_1//(R_2 + r_o//R_L)} v_S$$

or

$$a_{\text{ft}} = \frac{v_L}{v_S} = \frac{1}{1 + \dfrac{R_2}{r_o//R_L}} \times \frac{1}{1 + \dfrac{R_s + r_d}{R_1//(R_2 + r_o//R_L)}}$$

(*b*) Plugging in the given data we get

$$T = 10^3 \times \frac{1}{2.5} \times \frac{1}{10.36} \times \frac{1}{1.434} = 26.93 \qquad a_{\text{ft}} = \frac{1}{13} \times \frac{1}{28.56} = 2.693 \times 10^{-3} \text{ V/V}$$

Also, $A_{\text{ideal}} = 1 + R_2/R_1 = 1 + 9/1 = 10$ V/V, so Eq. (1.72) finally gives,

$$\frac{v_L}{v_S} = \frac{10}{1 + 1/26.93} + \frac{2.693 \times 10^{-3}}{1 + 26.93} = 9.642 + 9.64 \times 10^{-5} = 9.642 \text{ V/V}$$

Evidently a_{ft} has a negligible effect in this case. Note that in this example return-ratio analysis has taken automatically into account both R_s and R_L.

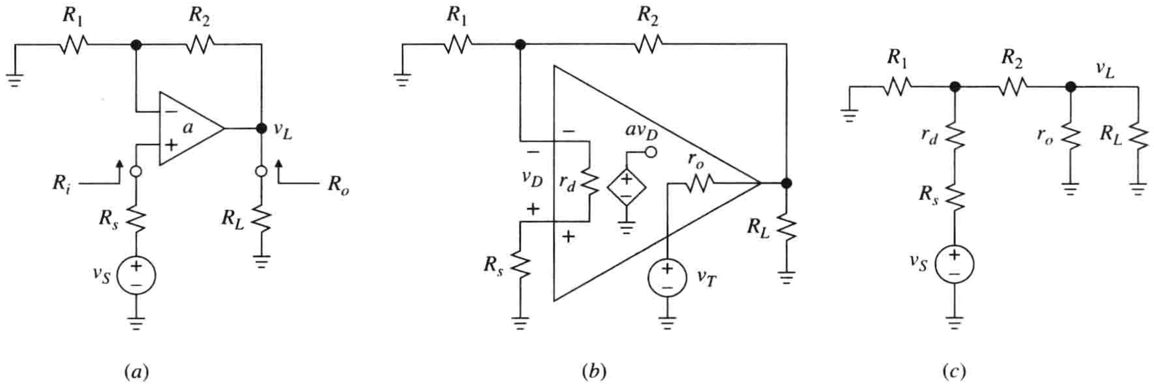


FIGURE 1.39
(*a*) Circuit of Example 1.12. Circuits to find (*b*) T and (*c*) a_{ft}.

Blackman's Impedance Formula

The return ratio provides a powerful tool also for the calculation of the *closed-loop resistance R* between any pair of nodes of a negative-feedback circuit—not just the nodes of the input and output ports. Such a resistance is found via the *Blackman's impedance formula*[6] as

$$R = r_0 \frac{1 + T_{\text{sc}}}{1 + T_{\text{oc}}} \tag{1.79a}$$

where

$$r_0 = \lim_{a \to 0} R \tag{1.79b}$$

is the resistance between the given node pair with the dependent source av_D set to zero, and T_{sc} and T_{oc} are the return ratios with the two nodes *short-circuited* and *open-circuited*, respectively. Blackman's formula holds regardless of the feedback topology in use.[2] Typically, either T_{sc} or T_{oc} is zero, indicating a series topology if $T_{\text{oc}} = 0$, and a shunt topology if $T_{\text{sc}} = 0$.

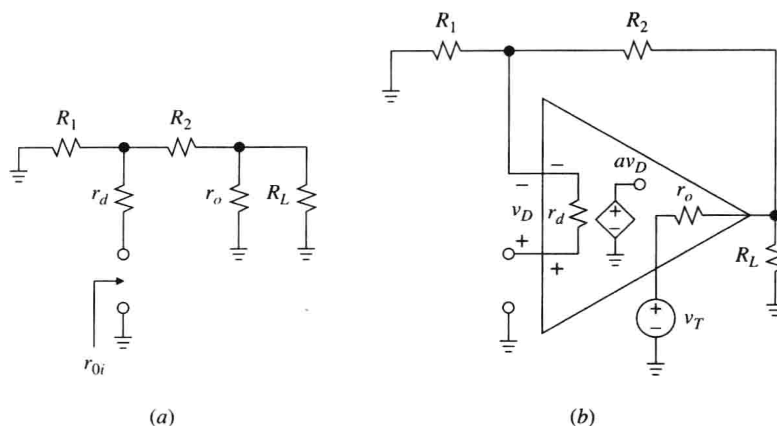


FIGURE 1.40
Circuits to find the noninverting amplifier's input resistance R_i via Blackman's formula.

EXAMPLE 1.13. For the noninverting amplifier of Example 1.12, use Blackman's formula to find (*a*) the resistance R_i presented to the input source and (*b*) the resistance R_o presented to the output load.

Solution.

(*a*) The input-port resistance with the dependent source set to zero is shown in Fig. 1.40*a*, where

$$r_{0i} = r_d + R_1 // (R_2 + r_o + R_L)$$

With the input port open-circuited as in Fig. 1.40*b*, no current flows through r_D, so $v_D = 0$ and thus $T_{oc} = 0$. Short-circuiting the input port turns the circuit into that of Fig. 1.39*b*, but with $R_s = 0$. Recalculating,

$$T_{sc} = a \times \frac{1}{1 + \dfrac{R_2}{r_d // R_1}} \times \frac{1}{1 + \dfrac{r_o}{(r_d // R_1 + R_2) // R_L}} \qquad R_i = r_{0i} \frac{1 + T_{sc}}{1 + 0} = r_{0i}(1 + T_{sc})$$

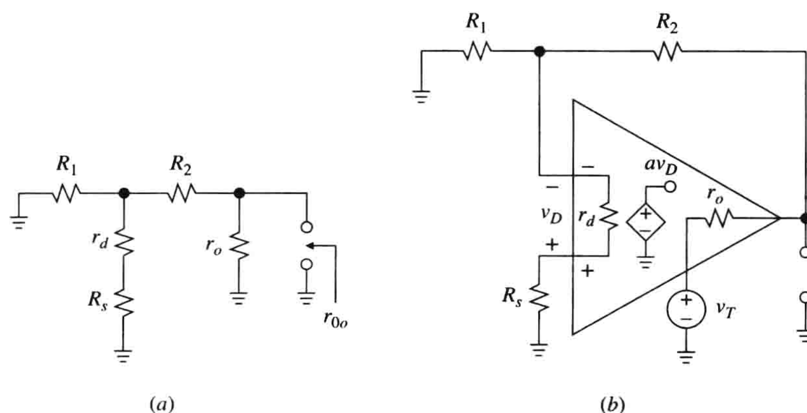


FIGURE 1.41
Circuits to find the noninverting amplifier's output resistance R_o via Blackman's formula.

(*b*) The output-port resistance with the dependent source set to zero is shown in Fig. 1.41*a*, where

$$r_{0o} = r_o//[R_2 + R_1//(r_d + R_s)]$$

With the output port open-circuited as in Fig. 1.41*b*, the circuit reduces to that of Fig. 1.39*b* but with $R_L = \infty$. Recycling, and using the fact that short-circuiting the output port gives $v_D = 0$ and thus $T_{sc} = 0$, we get

$$T_{oc} = a \times \frac{1}{1 + \dfrac{R_2}{r_d//R_1}} \times \frac{1}{1 + \dfrac{r_o}{r_d//R_1 + R_2}} \qquad R_o = r_{0o}\frac{1+0}{1+T_{OC}} = \frac{r_{0o}}{1+T_{OC}}$$

EXAMPLE 1.14. The inverting amplifier of Fig. 1.42*a* utilizes a T network in its feedback loop to achieve a high gain while using a relatively large resistance R_1 to ensure high input resistance. (*a*) Assuming an op amp with $a = 10^5$ V/V, $r_d = 1.0\ \text{M}\Omega$, and $r_o = 100\ \Omega$, find A_{ideal}, T, and A if $R_1 = R_2 = 1.0\,\text{M}\Omega$, $R_3 = 100\,\text{k}\Omega$, and $R_4 = 1.0\,\text{k}\Omega$. Can you ignore feedthrough in this case? (*b*) Use Blackman's formula to estimate the input and output resistances R_i and R_o. Verify with PSpice.

Solution.

(*a*) In the ideal case we have $v_N = 0$, and the voltage at the node common to R_2, R_3, and R_4 is, by the inverting-amplifier formula, $v_X = -(R_2/R_1)v_I$. Summing currents at v_X gives $(0 - v_X)/R_2 + (0 - v_X)/R_4 + (v_O - v_X)/R_3 = 0$. Eliminating v_X and solving for the ratio v_O/v_I gives

$$A_{\text{ideal}} = -\frac{R_2}{R_1}\left(1 + \frac{R_3}{R_2} + \frac{R_3}{R_4}\right) = -101.1\ \text{V/V}$$

Turning to Fig. 1.42*b*, we apply the voltage divider formula twice to write

$$v_D = -\frac{R_1//r_d}{R_1//r_d + R_2} \times \frac{[R_1//r_d + R_2]//R_4}{[R_1//r_d + R_2]//R_4 + R_3 + r_o} v_T = -\frac{v_T}{303.5}$$

so Eqs. (1.78) and (1.72) give

$$T = -\frac{10^5}{-303.5} = 329.5 \qquad A \cong \frac{A_{\text{ideal}}}{1 + 1/T} = \frac{-101.1}{1 + 1/329.5} = -100.8\ \text{V/V}$$

Feedthrough around the op amp takes place via the high-resistance path R_1–R_2, gets shunted by the comparatively low resistance R_4, and what's left continues via the high-resistance path R_3, to get shunted once again by the low resistance r_o, so we can indeed ignore the term $a_{\text{ft}}/(1 + T)$ in this example.

(*b*) To find R_i, note that the resistance presented by the input port in the limit $av_D \to 0$ is

$$r_{0i} = R_1 + r_d//[R_2 + R_4//(R_3 + r_o)] = 1.5002\ \text{M}\Omega$$

With the input port short-circuited, we have $T_{sc} = T = 329.5$. With the input port open-circuited, we have $T_{oc} = T(R_1 = \infty)$, so recalculating we get $T_{oc} = 494.3$. By Blackman's formula,

$$R_i = 1.5002\frac{1 + 329.5}{1 + 494.3} = 1.001\ \text{M}\Omega$$

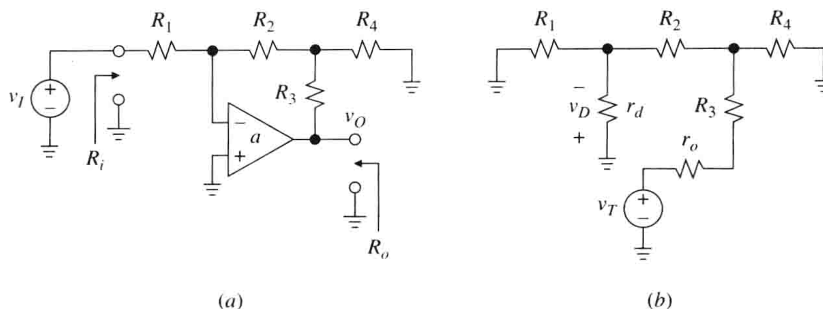


FIGURE 1.42
(*a*) Inverting amplifier with *T* network, and (*b*) circuit to find its return ratio $\Re$.

To find R_o, note that the resistance presented by the output port in the limit $av_D \rightarrow 0$ is

$$r_{0o} = r_o//(R_3 + \cdots) \cong r_o = 100\ \Omega$$

In this case $T_{sc} = 0$ and $T_{oc} = T = 329.5$, so

$$R_o = 100\frac{1+0}{1+329.5} = 0.302\ \Omega$$

A PSpice simulation confirms all the above calculations.

Comparing *T* and *L*

Though T and L may seem deceptively similar, they are generally different. Because of feedthrough, T and L contribute *differently* to the closed-loop gain A, as depicted in Figs. 1.37 and 1.34. Even when $T = L$, as it was the case in Example 1.11 because of the assumptions $r_d = \infty$ and $r_o = 0$, the contributions to A still differ because of a_{ft}. Compared to the analysis based on L, that based on T is more insightful because it splits A into two separate components, both stemming from forward transmission, but one through the error amplifier and the other through the feedback network. The analysis based on L, instead, lumps all forward transmission exclusively through the error amplifier so as to conform to the simpler diagram of Fig. 1.34. As such, it gives only approximate results compared to the exact results of return-ratio analysis,[2] though the difference may be very slight if T and L are sufficiently large.

It is apparent from Eq. (1.72) that if the condition

$$|\ a_{ft}\ | \ll |\ TA_{ideal}\ | \tag{1.80}$$

is met, we can ignore the feedthrough component in Eq. (1.72), after which Eq. (1.72) becomes formally identical to Eq. (1.68*a*), though T and L may still differ and thus give slightly different values of A. It is instructive to view the signal component $a_{ft}v_I$ that manages to creep to the output as a form of output noise. Reflected to the input, this noise gets divided by the gain TA_{ideal}, thus resulting in the equivalent input noise $a_{ft}v_I/(TA_{ideal})$. As long as Eq. (1.80) is met, this component will be

much smaller than v_I, so the fact that the feedback network is not unilateral is of little consequence in this case.

In Chapter 6 we will see that the gain a of an op amp rolls off with frequency, in turn causing both T and L to decrease at high frequencies. Moreover, because of reactive parasitics, r_d turns capacitive at high frequencies, and r_o may exhibit inductive behavior, further accentuating the difference between L and T, and also causing a_{ft} to increase with frequency, at least up to a point. Combine this with the frequency rolloff of T, and we have good reason to expect the component of A due to a_{ft} in Eq. (1.72) to become more relevant at high frequencies.

But, there are additional reasons for us to favor return-ratio analysis. The derivation of L in the examples of the previous section required that we first identify the feedback topology by examining the type of summing and sampling, and then that we derive a_ε and b *separately* to finally find $L = a_\varepsilon b$. By contrast, the derivation of T is a *one-shot affair*, regardless of the topology. (For topology identification we can always enlist Blackman's formula, which indicates a series type if T_{sc} prevails, a shunt type if T_{oc} prevails.) So, why bother with L at all? Fact is, there is a large body of literature, known as *two-port analysis*,[2,3] which derives L by dissecting a feedback circuit into *a feedback block* with b depending exclusively on the components of the feedback network, and an *amplifier block* with a_ε lumping together both amplification and loading by the feedback network. By contrast (we shall see in greater detail in Chapter 6), return-ratio analysis shifts loading from the amplifier to the feedback network to establish an alternative form of feedback factor that we shall denote as β to distinguish it from the factor b of two-port analysis. Although introduced by H. W. Bode[4] in the mid-nineteen forties, the return-ratio viewpoint has gained increasing prominence only more recently.[5,7,8]

1.8 OP AMP POWERING

In order to function, op amps need to be externally powered. Powering serves the twofold purpose of biasing the internal transistors and providing the power that the op amp must in turn supply to the output load and the feedback network. Figure 1.43 shows a recommended way of powering ops amps (though the figure shows the bipolar supplies V_{CC} and V_{EE}, the present considerations hold also for the CMOS supplies V_{DD} and V_{SS}). To prevent the ac noise usually present on the supply lines from interfering with the op amps, the supply pins of each IC must be bypassed to ground by means of low-inductance capacitors (0.1-μF ceramic capacitors are usually adequate). These decoupling capacitors also help neutralize any spurious feedback loops arising from the nonzero impedances of the supply and ground lines, or busses, which might pose stability problems. For this cure to be effective, the leads must be kept short to minimize their distributed inductance, which rises at the rate of about 1 nH/mm, and the capacitors must be mounted as close as possible to the op amp pins. A well-constructed circuit board will also include 10-μF polarized capacitors at the points of entry of the supply voltages to provide board-level bypass. Moreover, using wide ground traces will help maintain an electrically clean ground reference.

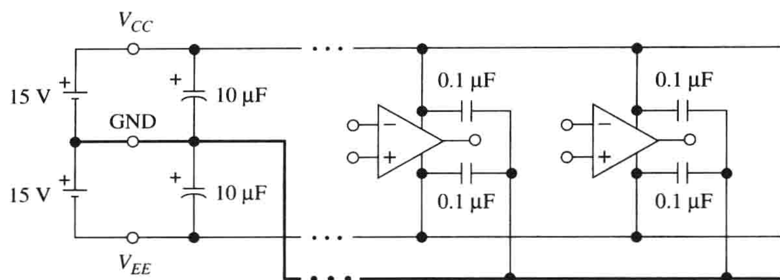


FIGURE 1.43
Op amp powering with bypass capacitors.

Typically, op amps are powered by means of a dual regulated power supply. Though the ±15-V values shown in Fig. 1.43 have long been the standard in analog systems, today's mixed-mode products incorporating both digital and analog functions on the same chip call for a single, lower-valued supply such as $V_{CC} = 5$ V and $V_{EE} = 0$, or $V_{DD} = 3.3$ V and $V_{SS} = 0$, or, in more recent devices, $V_{DD} = 0.8$ V and $V_{SS} = 0$. To reduce cluttering, the power supplies are normally omitted from circuit diagrams.

Current Flow and Power Dissipation

Since virtually no current flows in or out of the input pins of an op amp, the only current-carrying terminals are the output and the supply pins. We shall designate their currents as i_O, i_{CC}, and i_{EE}. Since V_{CC} is the most positive and V_{EE} the most negative voltage in the circuit, under proper operation i_{CC} will always flow *into* and i_{EE} always *out of* the op amp. However, i_O may flow either out of or into the op amp, depending on circuit conditions. In the former case the op amp is said to be *sourcing* current, and in the latter it is *sinking* current. At all times, the three currents must satisfy KCL. So for an op amp sourcing current we have $i_{CC} = i_{EE} + i_O$, and for an op amp sinking current we have $i_{EE} = i_{CC} + i_O$.

In the special case in which $i_O = 0$, we have $i_{CC} = i_{EE} = I_Q$, where I_Q is called the *quiescent supply current.* This is the current that biases the internal transistors to keep them electrically alive. Its magnitude depends on the op amp type and, to a certain extent, on the supply voltages; typically, I_Q is in the milliampere range. Op amps intended for portable equipment applications may have I_Q in the microampere range and are therefore called *micropower op amps.*

Figure 1.44 shows current flow in the noninverting and inverting circuits, both for the case of a positive and a negative input. Trace each circuit in detail until you are fully convinced that the various currents flow as shown. Note that the output current consists of two components, one to feed the load and the other to feed the feedback network. Moreover, the flow of currents I_Q and i_O through the op amp causes *internal power dissipation*. This dissipation must never exceed the maximum rating specified in the data sheets.

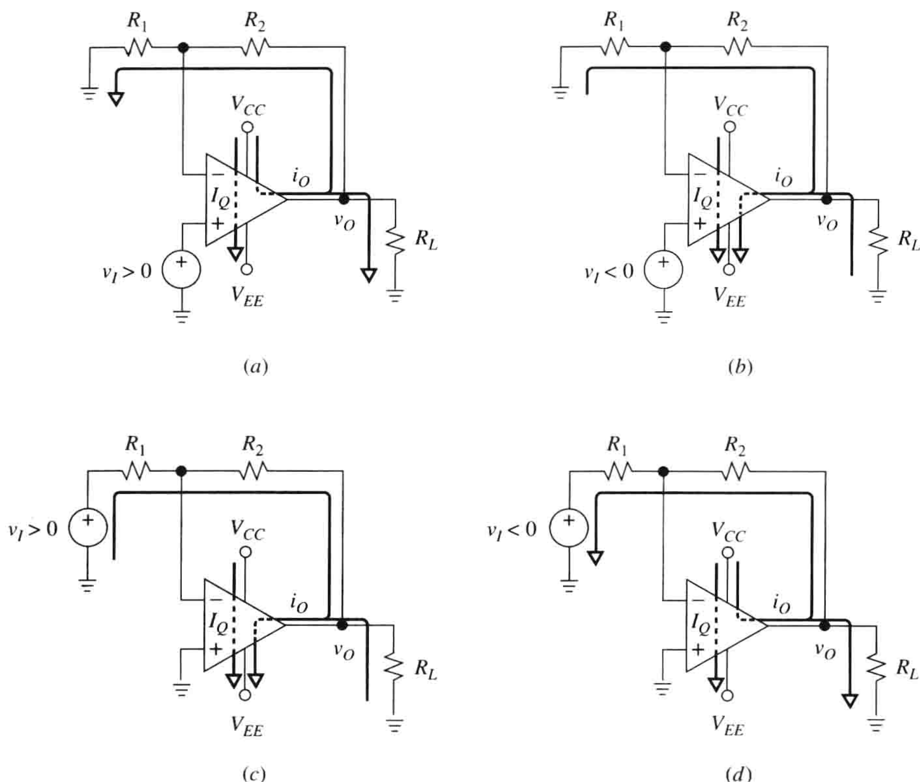


FIGURE 1.44
Current flow for the noninverting [(*a*) and (*b*)] and the inverting [(*c*) and (*d*)] amplifiers.

EXAMPLE 1.15. An inverting amplifier with $R_1 = 10\ \text{k}\Omega$, $R_2 = 20\ \text{k}\Omega$, and $v_I = 3$ V drives a 2-kΩ load. (*a*) Assuming $\pm$15-V supplies and $I_Q = 0.5$ mA, find i_{CC}, i_{EE}, and i_O. (*b*) Find the power dissipated inside the op amp.

Solution.

(*a*) With reference to Fig. 1.44*c*, we have $v_O = -(20/10)3 = -6$ V. Denoting the currents through R_L, R_2, and R_1 as i_L, i_2, and i_1, we have $i_L = 6/2 = 3$ mA, and $i_2 = i_1 = 3/10 = 0.3$ mA. Thus, $i_O = i_2 + i_L = 0.3 + 3 = 3.3$ mA; $i_{CC} = I_Q = 0.5$ mA; $i_{EE} = i_{CC} + i_O = 0.5 + 3.3 = 3.8$ mA.

(*b*) Whenever a current i experiences a voltage drop v, the corresponding power is $p = vi$. Thus, $p_{OA} = (V_{CC} - V_{EE})I_Q + (v_O - V_{EE})i_O = 30 \times 0.5 + [-6 - (-15)] \times 3.3 = 44.7$ mW.

EXAMPLE 1.16. When experimenting with op amps, it is handy to have a variable source over the range $-10\ \text{V} \le v_S \le 10$ V. (*a*) Design one such source using a 741 op amp and a 100-kΩ potentiometer. (*b*) If v_S is set to 10 V, how much does it change when we connect a 1-kΩ load to the source?

Solution.

(*a*) We first design a resistive network to produce an adjustable voltage over the range -10 V to $+10$ V. As shown in Fig. 1.45, where we use a concise notation for the supply voltages, this network consists of the potentiometer and two 25-kΩ resistors

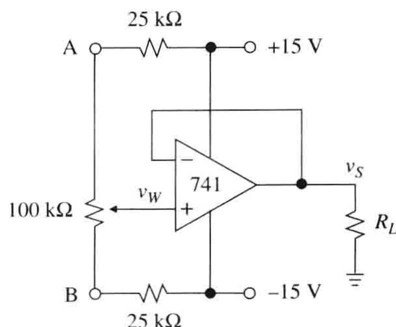


FIGURE 1.45
Variable source from -10 V to $+10$ V.

to drop 5 V each, so that $v_A = 10$ V and $v_B = -10$ V. By turning the wiper, we can vary v_W over the range $-10\text{ V} \le v_W \le 10$ V. However, if a load is connected directly to the wiper, v_W will change significantly because of the loading effect. For this reason we interpose a unity-gain buffer, as shown.

(*b*) Connecting a 1-kΩ load will draw a current $i_L = 10/1 = 10$ mA. The output resistance is $R_o = r_o/(1+T) = 75/(1+200{,}000) = 0.375$ mΩ. The source change is thus $\Delta v_S = R_o \Delta i_L = 0.375 \times 10^{-3} \times 10 \times 10^{-3} = 3.75\ \mu$V—quite a small change! This demonstrates a most important op amp application, namely, *regulation* against changes in the load conditions.

Output Saturation

The supply voltages V_{CC} and V_{EE} set upper and lower bounds on the output swing capability of the op amp. This is best visualized in terms of the VTC of Fig. 1.46, which reveals three different regions of operation.

In the *linear region* the curve is approximately straight and its slope represents the open-loop gain a. With a as large as the 741s 200,000 V/V, the curve is so steep that it practically coalesces with the vertical axis, unless we use different scales for the two axes. If we express v_O in volts and v_D in microvolts, as shown, then the slope becomes 0.2 V/μV. As we know, op amp behavior within this region is modeled with a *dependent* source of value av_D.

As v_D is increased, v_O increases in proportion until a point is reached where internal transistor saturation effects take place that cause the VTC to flatten out. This is the *positive saturation region,* where v_O no longer depends on v_D but remains fixed, making the op amp behave as an *independent* source of value V_{OH}. Similar considerations hold for the *negative saturation region,* where the op amp acts as an independent source of value V_{OL}. Note that in saturation v_D is no longer necessarily in the microvolt range!

The 741 data sheets, reproduced at the end of Chapter 5, indicate that with $\pm$15-V supplies and with a typical output load of 2 kΩ, the 741 saturates at $\pm V_{\text{sat}} \cong \pm 13$ V, that is, within 2 V of the supply rails. The *output voltage swing*, defined as $\text{OVS} = V_{OH} - V_{OL}$, is in this case $\text{OVS} \cong 13 - (-13) = 26$ V, also expressed as $\text{OVS} \cong \pm 13$ V. Moreover, since $13/200{,}000 = 65\ \mu$V, the input voltage range corresponding to the linear region is thus $-65\ \mu\text{V} \le v_D \le +65\ \mu$V.

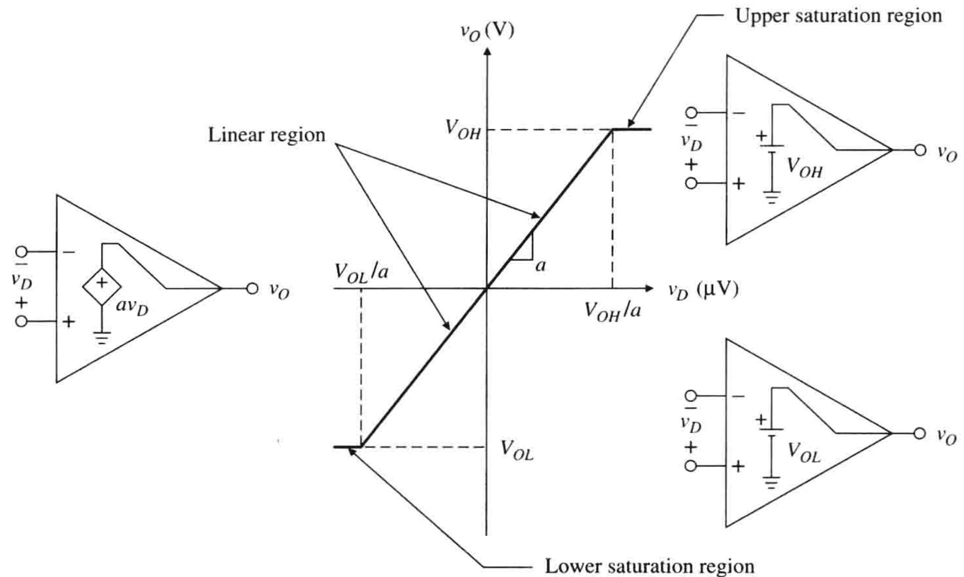


FIGURE 1.46
Regions of operation and approximate op amp models.

If the supplies are other than ±15 V, the 741 saturation voltages will change accordingly. For instance, a 741 powered from a single 9-V battery and driving a 2-kΩ load can be assumed to saturate at $V_{OH} \cong 9-2=7$ V and $V_{OL} \cong 0+2=2$ V, so now OVS $\cong 7-2=5$ V.

In single-supply systems, such as mixed digital-analog systems with $V_{CC} = 5$ V and $V_{EE} = 0$ V, signals are usually constrained within the range of 0 V to 5 V. The need arises for a reference voltage at $(1/2)V_{CC} = 2.5$ V for termination of all analog sources and loads, and thus allow for symmetric voltage swings about this common reference. In Fig. 1.47 this voltage is synthesized by the R-R voltage divider, and is then buffered by OA_1 to provide a low-resistance drive. To maximize the dynamic range of signals, OA_2 is typically a device with rail-to-rail output capabilities, or $V_{OH} \cong 5$ V and $V_{OL} \cong 0$ V. The TLE2426 Rail Splitter is a 3-terminal chip containing

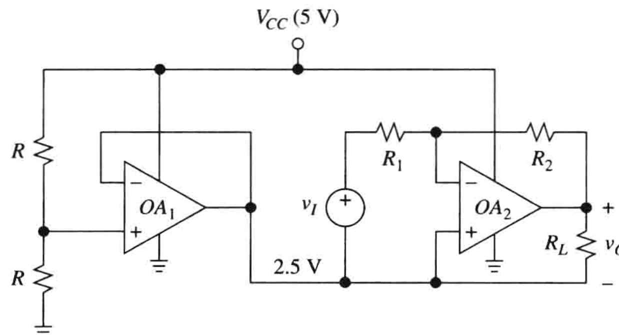


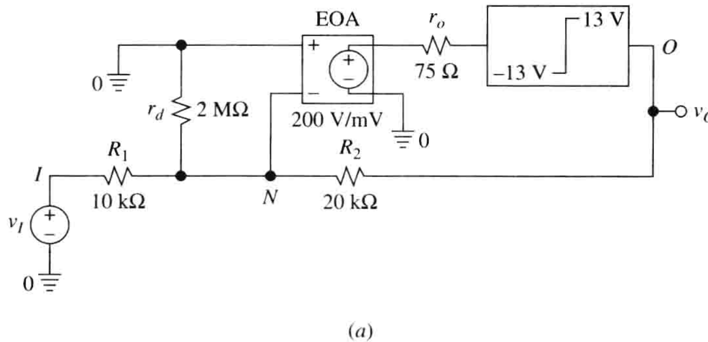
FIGURE 1.47
Synthesis of a 2.5-V common reference voltage in a 5-V single-supply system.

all circuitry needed for the synthesis of a precision 2.5-V common reference with a 7.5-mΩ output resistance.

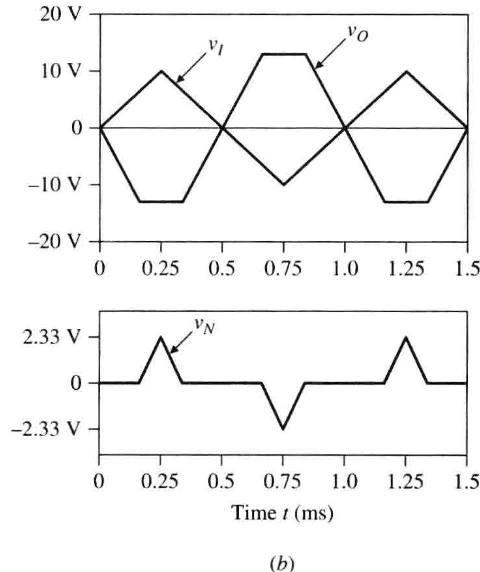
When an op amp is used in the *negative-feedback mode,* its operation must be confined within the linear region because only there is the op amp capable of influencing its own input. If the device is inadvertently pushed into saturation, v_O will remain fixed and the op amp will no longer be able to influence v_D, thus resulting in completely different behavior.

SPICE Simulation

The basic 741 SPICE model of Fig. 1.5 does not saturate. To simulate saturation, we use the limiter block available in the PSpice library. The PSpice circuit of Fig. 1.48*a*



(*a*)



(*b*)

FIGURE 1.48
(*a*) Inverting amplifier and (*b*) waveforms when the op amp is driven in saturation.

uses a ± 13 V limiter to simulate a 741 op amp operating as an inverting amplifier with a gain of $A = -20/10 = -2$ V/V.

If we drive the circuit with a triangular wave of, say, ± 5-V peak values, the output will be an (inverted) triangular wave of ∓ 10-V peak values, which are well within the permissible output range of ± 13 V. Moreover, the inverting input $v_N = -v_O/a$ will also be a triangular wave of $\pm 10/200{,}000 = \pm 50$-$\mu$V peak values, small enough for v_N to be considered fairly close to a virtual ground. However, if we overdrive the circuit with a triangular wave of ± 10-V peak values as shown, we cannot expect it to yield an (inverted) wave of ∓ 20-V peak values because they exceed the permissible output range. The circuit will give $v_O = -2v_I$ only so long as the op amp is allowed to operate within its linear region, $-13 \text{ V} \le v_O \le +13$ V, in turn corresponding to $-6.5 \text{ V} \le v_I \le +6.5$ V. As soon as v_I exceeds this range, v_O will saturate at ± 13 V, resulting in the *clipped* waveform shown in Fig. 1.48*b*.

It is instructive to observe also the waveform v_N, which is close to a virtual ground only so long as the op amp is within its linear region. Once driven in saturation, the op amp is deprived of its ability to influence v_N via the feedback network, so v_N will no longer approximate a virtual ground. In saturation we have, by the superposition principle,

$$v_{N(\text{sat})} = \frac{R_2}{R_1 + R_2} v_I + \frac{R_1}{R_1 + R_2}(\pm V_{\text{sat}}) = \frac{2}{3} v_I + \frac{1}{3}(\pm 13) = \frac{2v_I \pm 13 \text{ V}}{3}$$

In particular, when v_I peaks at 10 V, $v_{N(\text{sat})}$ peaks at $(2 \times 10 - 13)/3 = 2.33$ V, as shown in the figure (by symmetry, $v_{N(\text{sat})}$ peaks at -2.33 V when v_I peaks at -10 V).

Clipping is a form of distortion because the output waveform of a linear amplifier is expected to have the same shape as the input. Clipping is generally undesirable, though there are situations in which it is exploited on purpose to achieve specific effects. To avoid clipping, we must either keep v_I below a suitable level, or suitably lower the amplifier's gain A.

PROBLEMS

1.1 Amplifier fundamentals

1.1 A voltage amplifier of the type of Fig. 1.1 is fed by a 200-mV source and drives a 10-Ω load. Voltmeter measurements give $v_I = 150$ mV and $v_O = 10$ V. If disconnecting the load causes v_O to rise to 12 V, and then connecting a 30-kΩ resistor across the input-port terminals causes v_O to drop from 12 V to 9.6 V, find R_s, R_i, A_{oc}, and R_o.

1.2 A source v_S with a 1-kΩ internal resistance is fed to a current amplifier of the type of Fig. 1.2 having $R_i = 200\ \Omega$, $A_{\text{sc}} = 180$ A/A, and $R_o = 10$ kΩ. The amplifier in turn drives a load $R_L = 2$ kΩ, whose voltage is denoted as v_L. Sketch and label the circuit, and find the voltage gain v_L/v_S as well as the power gain p_L/p_S, where p_S is the power released by the source v_S and p_L is that absorbed by the load R_L.

1.3 Sketch and label a transresistance amplifier being driven by a source $i_S = 3\ \mu$A with internal resistance $R_s = 100$ kΩ. Given that the voltage across the input port is 50 mV,

and the output-port open-circuit voltage and short-circuit current are, respectively, 10 V and 20 mA, find R_i, A_{oc}, and R_o, as well as the gain v_O/i_S, if the amplifier is terminated on a 1.5-kΩ load.

1.4 Sketch and label a transconductance amplifier that is driven by a source $v_S = 150$ mV having a 100-kΩ internal resistance, and drives a load R_L. Given that $v_I = 125$ mV, $i_L = 230$ mA for $R_L = 40\ \Omega$, and $i_L = 220$ mA for $R_L = 60\ \Omega$, find R_i, A_{oc} and R_o. What is its short-circuit output current with $v_S = 100$ mV?

1.5 A transresistance amplifier with $R_{i1} = 100\ \Omega$, $A_{oc1} = 0.2$ V/mA, and $R_{o1} = 100\ \Omega$, is driven by a source i_S having internal resistance $R_s = 1$ kΩ, and drives in turn a transconductance amplifier having $R_{i2} = 1$ kΩ, $A_{sc2} = 100$ mA/V, and $R_{o2} = 100$ kΩ. The latter, in turn, drives a 25-kΩ load. Sketch and label the circuit, and find the source-to-load current gain i_L/i_S as well as the power gain p_L/p_S, where p_S is the power released by the source v_S and p_L is that absorbed by the load R_L.

1.2 The operational amplifier

1.6 Given an op amp with $r_d \cong \infty$, $a = 10^4$ V/V, and $r_o \cong 0$, find (*a*) v_O if $v_P = 750.25$ mV and $v_N = 751.50$ mV, (*b*) v_N if $v_O = -5$ V and $v_P = 0$, (*c*) v_P if $v_N = v_O = 5$ V, and (*d*) v_N if $v_P = -v_O = 1$ V.

1.7 An op amp with $r_d = 1$ MΩ, $a = 100$ V/mV, and $r_o = 100\ \Omega$ is driving a 2-kΩ load and is part of a circuit such that $v_P = -2.0$ mV and $v_O = -10$ V. Sketch the circuit, and find the voltages across and the currents through r_d and r_o (make sure you indicate voltage polarities and current directions). What is the value of v_N?

1.3 Basic op amp configurations

1.8 (*a*) Find the minimum value of a needed to ensure that the gain of the voltage follower of Fig. 1.8*a* does not depart from +1.0 V/V by more than 0.01%. (*b*) Repeat, but for the case of the inverting amplifier of Fig. 1.10*a* configured for a gain of −1.0 V/V with $R_1 = R_2$. Why the difference?

1.9 (*a*) Design a noninverting amplifier whose gain is variable over the range 1 V/V $\leq A \leq$ 5 V/V by means of a 100-kΩ pot. (*b*) Repeat (*a*) for 0.5 V/V $\leq A \leq$ 2 V/V. *Hint:* To achieve $A \leq 1$ V/V, you need an input voltage divider.

1.10 (*a*) A noninverting amplifier is implemented with two 10-kΩ resistances having 5% tolerance. What is the range of possible values for the gain A? How would you modify the circuit for the exact calibration of A? (*b*) Repeat, but for the inverting amplifier.

1.11 In the inverting amplifier of Fig. 1.10*a*, let $v_I = 0.1$ V, $R_1 = 10$ kΩ, and $R_2 = 100$ kΩ. Find v_O and v_N if (*a*) $a = 10^2$ V/V, (*b*) $a = 10^4$ V/V, and (*c*) $a = 10^6$ V/V. Comment on your findings.

1.12 (*a*) Design an inverting amplifier whose gain is variable over the range −10 V/V $\leq A \leq 0$ by means of a 100-kΩ pot. (*b*) Repeat, but for −10 V/V $\leq A \leq -1$ V/V. *Hint:* To prevent A from reaching zero, you must use a suitable resistor in series with the pot.

1.13 (*a*) A source $v_S = 2$ V with $R_s = 10$ kΩ is to drive a gain-of-five inverting amplifier implemented with $R_1 = 20$ kΩ and $R_2 = 100$ kΩ. Find the amplifier output voltage and verify that because of loading its magnitude is *less* than $2 \times 5 = 10$ V. (*b*) Find the value to which R_2 must be changed if we want to compensate for loading and obtain a full output magnitude of 10 V.

1.14 (*a*) A source $v_S = 10$ V is fed to a voltage divider implemented with $R_A = 120$ kΩ and $R_B = 30$ kΩ, and the voltage across R_B is fed, in turn, to a gain-of-five noninverting amplifier having $R_1 = 30$ kΩ and $R_2 = 120$ kΩ. Sketch the circuit, and predict the amplifier output voltage v_O. (*b*) Repeat (*a*) for a gain-of-five inverting amplifier having $R_1 = 30$ kΩ and $R_2 = 150$ kΩ. Compare and comment on the differences.

1.4 Ideal op amp circuit analysis

1.15 Find v_N, v_P, and v_O in the circuit of Fig. P1.15, as well as the power released by the 4-V source; devise a method to check your results.

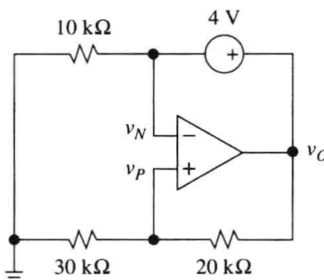


FIGURE P1.15

1.16 (*a*) Find v_N, v_P, and v_O in the circuit of Fig. P1.16. (*b*) Repeat (*a*) with a 5-kΩ resistance connected between A and B.

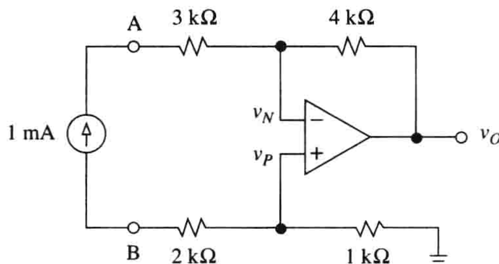


FIGURE P1.16

1.17 (*a*) Find v_N, v_P, and v_O in the circuit of Fig. P1.17 if $v_S = 9$ V. (*b*) Find the resistance R that, if connected between the inverting-input pin of the op amp and ground, causes v_O to double. Verify with PSpice.

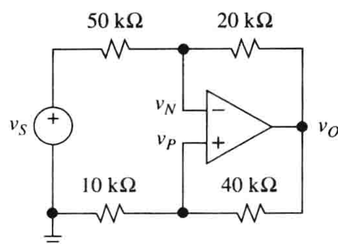


FIGURE P1.17

1.18 (*a*) Find v_N, v_P, and v_O in the circuit of Fig. P1.18. (*b*) Repeat (*a*) with a 40-kΩ resistance in parallel with the 0.3-mA source.

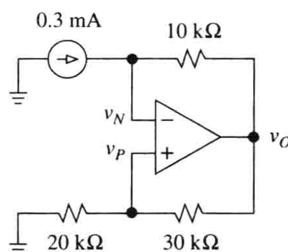


FIGURE P1.18

1.19 (*a*) Find v_N, v_P, and v_O in the circuit of Fig. P1.19 if $i_S = 1$ mA. (*b*) Find a resistance R that when connected in parallel with the 1-mA source will cause v_O to drop to half the value found in (*a*).

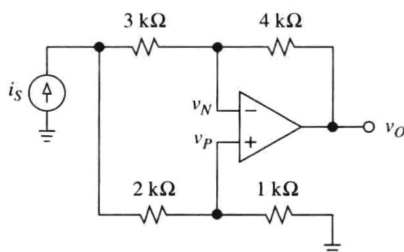


FIGURE P1.19

1.20 (*a*) Find the equivalent resistance seen by the current source in Fig. P1.16. *Hint:* Find the voltage across the source, and then obtain the resistance as the ratio of voltage over current. (*b*) Repeat for Fig. P1.18. (*c*) Repeat for Fig. P1.19.

1.21 (*a*) If the current source of Fig. P1.16 is replaced by a voltage source v_S, find the magnitude and polarity of v_S so that $v_O = 10$ V. (*b*) If the wire connecting the 4-V source to node v_O in Fig. P1.15 is cut and a 5-kΩ resistance is inserted in series between the two, to what value must the source be changed to yield $v_O = 10$ V?

1.22 In the circuit of Fig. P1.22 the switch is designed to provide gain-polarity control. (*a*) Verify that $A = +1$ V/V when the switch is open, and $A = -R_2/R_1$ when the

switch is closed, so that making $R_1 = R_2$ yields $A = \pm 1$ V/V. (*b*) To accommodate gains greater than unity, connect an additional resistance R_4 from the inverting-input pin of the op amp to ground. Derive separate expressions for A in terms of R_1 through R_4 with the switch open and with the switch closed. (*c*) Specify resistance values suitable for achieving $A = \pm 2$ V/V.

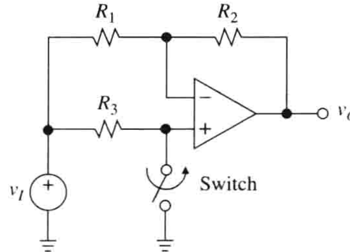


FIGURE P1.22

1.23 (*a*) Find the equivalent resistance seen by the voltage source of Fig. P1.15. *Hint:* Find the current through the source, and then obtain the resistance as the ratio of voltage over current. (*b*) Repeat for Fig. P1.17. (*c*) Repeat for Fig. P1.22 (consider both the case of the switch open and closed).

1.24 In the circuit of Fig. P1.24 the pot is used to control gain magnitude as well as polarity. (*a*) Letting k denote the fraction of R_3 between the wiper and ground, show that varying the wiper from bottom to top varies the gain over the range $-R_2/R_1 \le A \le 1$ V/V, so that making $R_1 = R_2$ yields -1 V/V $\le A \le +1$ V/V. (*b*) To accommodate gains greater than unity, connect an additional resistance R_4 from the op amp's inverting-input pin to ground. Derive an expression for A in terms of R_1, R_2, R_4, and k. (*c*) Specify resistance values suitable for achieving -5 V/V $\le A \le +5$ V/V.

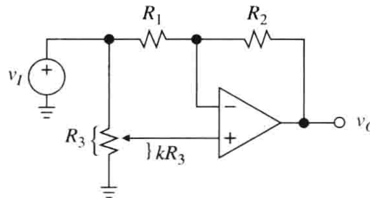


FIGURE P1.24

1.25 Consider the following statements about the input resistance R_i of the noninverting amplifier of Fig. 1.15*a*: (*a*) Since we are looking straight into the noninverting-input pin, which is an open circuit, we have $R_i = \infty$; (*b*) since the input pins are virtually shorted together, we have $R_i = 0 + (R_1 \parallel R_2) = R_1 \parallel R_2$; (*c*) since the noninverting-input pin is virtually shorted to the inverting-input pin, which is in turn a virtual-ground node, we have $R_i = 0 + 0 = 0$. Which statement is correct? How would you refute the other two?

1.26 (*a*) Show that the circuit of Fig. P1.26 has $R_i = \infty$ and $A = -(1 + R_3/R_4)R_1/R_2$. (*b*) Specify suitable components to make A variable over the range -100 V/V $\le A \le 0$ by means of a 100-kΩ pot. Try minimizing the number of resistors you use.

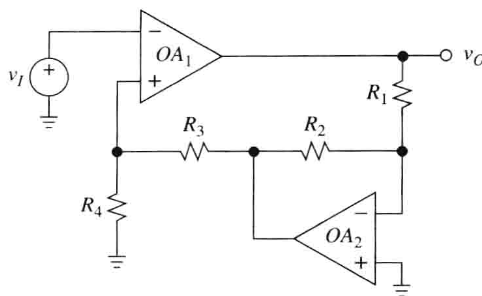


FIGURE P1.26

1.27 The audio panpot circuit of Fig. P1.27 is used to continuously vary the position of signal v_I between the left and the right stereo channels. (*a*) Discuss circuit operation. (*b*) Specify R_1 and R_2 so that $v_L/v_I = -1$ V/V when the wiper is fully down, $v_R/v_I = -1$ V/V when the wiper is fully up, and $v_L/v_I = v_R/v_I = -1/\sqrt{2}$ when the wiper is halfway.

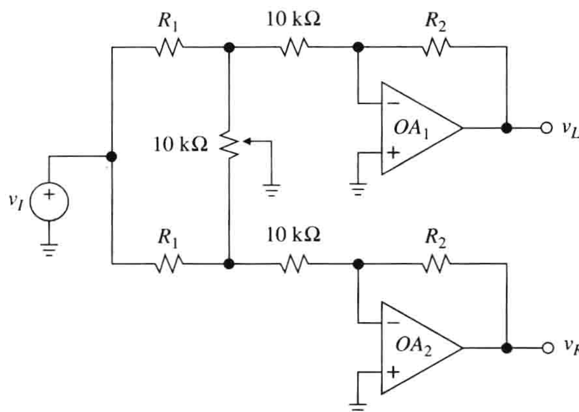


FIGURE P1.27

1.28 (*a*) Using standard 5% resistances in the kilohm range, design a circuit to yield $v_O = -100(4v_1 + 3v_2 + 2v_3 + v_4)$. (*b*) If $v_1 = 20$ mV, $v_2 = -50$ mV, and $v_4 = 100$ mV, find v_3 for $v_O = 0$ V.

1.29 (*a*) Using standard 5% resistances, design a circuit to give (*a*) $v_O = -10(v_I + 1\text{ V})$; (*b*) $v_O = -v_I + V_O$, where V_O is variable over the range $-5\text{ V} \le V_O \le +5\text{ V}$ by means of a 100-kΩ pot. *Hint:* Connect the pot between the ±15-V supplies and use the wiper voltage as one of the inputs to your circuit.

1.30 In the circuit of Fig. 1.18 let $R_1 = R_3 = R_4 = 10\text{ k}\Omega$ and $R_2 = 30\text{ k}\Omega$. (*a*) If $v_1 = 3$ V, find v_2 for $v_O = 10$ V. (*b*) If $v_2 = 6$ V, find v_1 for $v_O = 0$ V. (*c*) If $v_1 = 1$ V, find the range of values for v_2 for which $-10\text{ V} \le v_O \le +10\text{ V}$.

1.31 You can readily verify that if we put the output in the form $v_O = A_2v_2 - A_1v_1$ in the circuit of Fig. 1.18, then $A_2 \le A_1 + 1$. Applications requiring $A_2 \ge A_1 + 1$ can be accommodated by connecting an additional resistance R_5 from the node common to R_1 and R_2 to ground. (*a*) Sketch the modified circuit and derive a relationship between its

output and inputs. (*b*) Specify standard resistances to achieve $v_O = 5(2v_2 - v_1)$. Try to minimize the number of resistors you use.

1.32 (*a*) In the difference amplifier of Fig. 1.18 let $R_1 = R_3 = 10$ kΩ and $R_2 = R_4 = 100$ kΩ. Find v_O if $v_1 = 10\cos(2\pi 60t) - 0.5\cos(2\pi 10^3 t)$ V, and $v_2 = 10\cos(2\pi 60t) + 0.5\cos(2\pi 10^3 t)$ V. (*b*) Repeat if R_4 is changed to 101 kΩ. Comment on your findings.

1.33 Show that if all resistances in Fig. P1.33 are equal, then $v_O = v_2 + v_4 + v_6 - v_1 - v_3 - v_5$.

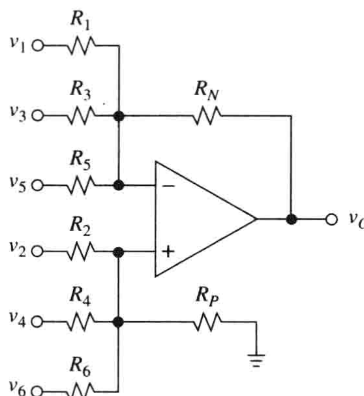


FIGURE P1.33

1.34 Using a topology of the type of Fig. P1.33, design a four-input amplifier such that $v_O = 4v_A - 3v_B + 2v_C - v_D$. Try minimizing the number of resistors you use.

1.35 Using just one op amp powered from ±12-V regulated supplies, design a circuit to yield: (*a*) $v_O = 10v_I + 5$ V; (*b*) $v_O = 10(v_2 - v_1) - 5$ V.

1.36 Using just one op amp powered from ±15-V supply voltages, design a circuit that accepts an ac input v_i and yields $v_O = v_i + 5$ V, under the constraint that the resistance seen by the ac source be 100 kΩ.

1.37 Design a two-input, two-output circuit that yields the sum and the difference of its inputs: $v_S = v_{I1} + v_{I2}$, and $v_D = v_{I1} - v_{I2}$. Try minimizing the component count.

1.38 Obtain a relationship between v_O and v_I if the differentiator of Fig. 1.19 includes also a resistance R_s in series with C. Discuss the extreme cases of v_I changing very slowly and very rapidly.

1.39 Obtain a relationship between v_O and v_I if the integrator of Fig. 1.20 includes also a resistance R_p in parallel with C. Discuss the extreme cases of v_I changing very rapidly and very slowly.

1.40 In the differentiator of Fig. 1.19 let $C = 10$ nF and $R = 100$ kΩ, and let v_I be a periodic signal alternating between 0 V and 2 V with a frequency of 100 Hz. Sketch and label v_I and v_O versus time if v_I is (*a*) a sine wave; (*b*) a triangular wave.

1.41 In the integrator of Fig. 1.20 let $R = 100$ kΩ and $C = 10$ nF. Sketch and label $v_I(t)$ and $v_O(t)$ if (*a*) $v_I = 5\sin(2\pi 100t)$ V and $v_O(0) = 0$; (*b*) $v_I = 5[u(t) - u(t - 2\text{ ms})]$ V

and $v_O(0) = 5$ V, where $u(t - t_0)$ is the unit step function defined as $u = 0$ for $t < t_0$, and $u = 1$ for $t > t_0$.

1.42 Suppose the integrator of Fig. 1.20 has $R = 100$ kΩ and $C = 10$ nF, and includes also a 300-kΩ resistance in parallel with C. (*a*) Assuming C is initially discharged, sketch and label $v_O(t)$ if v_I is changed from 0 V to +1 V at $t = 0$. *Hint:* The voltage across the capacitor can be expressed as $v(t \geq 0) = v_\infty + (v_0 - v_\infty)\exp[-t/(R_{eq}C)]$, where v_0 is the initial voltage, v_∞ is the steady-state value to which v tends in the limit $t \to \infty$ (when C acts as an open circuit), and R_{eq} is the equivalent resistance seen by C during the transient (to find R_{eq}, use the test method). (*b*) Sketch and label $v_O(t)$ if v_I is changed to -0.5 V at the instant in which v_O reaches -2 V.

1.43 Suppose the circuit of Fig. 1.42*a* uses 20-kΩ resistances throughout, and includes also a capacitance $C = 10$ nF between the output node and the inverting input node of the op amp. Assuming C is initially discharged, sketch and label $v_O(t)$ if v_I is changed from 0 V to +1 V at $t = 0$ (use the hint of Problem 1.42).

1.44 Show that if the op amp of Fig. 1.21*b* has a finite gain a, then $R_{eq} = (-R_1R/R_2)\times [1 + (1 + R_2/R_1)/a]/[1 - (1 + R_1/R_2)/a]$.

1.45 Find an expression for R_i in Fig. P1.45; discuss its behavior as R is varied over the range $0 \leq R \leq 2R_1$.

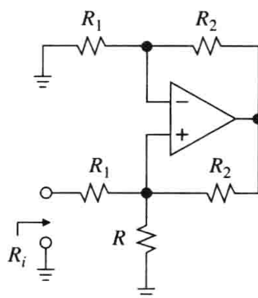


FIGURE P1.45

1.46 The circuit of Fig. P1.46 can be used to control the input resistance of the inverting amplifier based on OA_1. (*a*) Show that $R_i = R_1/(1 - R_1/R_3)$. (*b*) Specify resistances suitable for achieving $A = -10$ V/V with $R_i = \infty$.

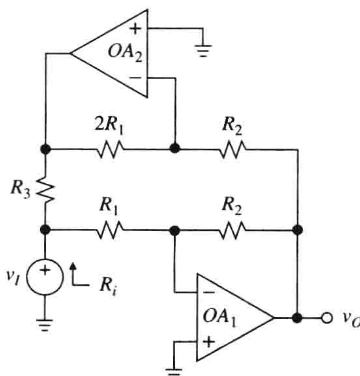


FIGURE P1.46

1.5 Negative feedback

1.47 (*a*) Find the feeback factor b of a negative-feedback system with $a_\varepsilon = 10^3$ and $A = 10^2$. (*b*) Find A exactly via Eq. (1.40), and approximately via Eq. (1.49) if a_ε drops by 10%. (*c*) Repeat (*b*) for a 50% drop in a_ε; compare with (*b*) and comment.

1.48 You are asked to design an amplifier with a gain A of 10^2 V/V that is accurate to within ±0.1%, or $A = 10^2$ V/V ± 0.1%. All you have available are amplifier stages with $a_\varepsilon = 10^4$ V/V ± 25% each. Your amplifier can be implemented using a cascade of basic stages, each employing a suitable amount of negative feedback. What is the minimum number of stages required? What is the feedback factor b of each stage?

1.49 A certain op amp has been configured as an inverting amplifier with two identical 10-kΩ resistors, and is driven by a 1-kHz sinusoidal wave having ±5-V peak values. Unfortunately, because of a fabrication error, the device exhibits $a = 10$ V/mV for $v_O > 0$, but only $a = 2.5$ V/mV for $v_O < 0$. Sketch and label v_I, v_O, and v_N versus time, and explain whether the device can still be used for the intended application.

1.50 The open-loop VTC of a certain op amp has a slope of 5 V/mV for $|v_O| \leq 2$ V, and 2 V/mV for $|v_O| > 2$ V. The op amp is used as a unity-gain voltage buffer, and is driven by a triangular wave with peak values of +4 V and −1 V, respectively. Sketch and label the open-loop VTC as well as the plots of v_I, v_O, and v_D versus time.

1.51 A crude BJT power amplifier of the class B (push-pull) type exhibits the VTC of Fig. P1.51*b*. The dead band occurring for $-0.7 \text{ V} \leq v_1 \leq +0.7$ V causes a crossover distortion at the output that can be reduced by preceding the power stage with a preamplifier stage and then using negative feedback to reduce the dead band. This is shown in Fig. P1.51*a* for the case of a difference preamplifier with gain a_1 and $b = 1$ V/V. (*a*) Sketch and label the closed-loop VTC if $a_1 = 10^2$ V/V. (*b*) Sketch v_I, v_1, and v_O versus time if v_I is a 100-Hz triangular wave with peak values of ±1 V.

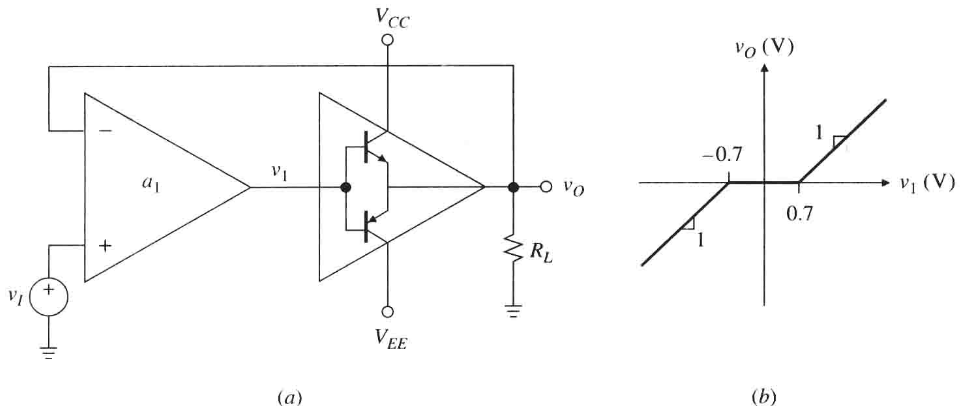


FIGURE P1.51

1.52 A certain audio power amplifier with a signal gain of 10 V/V is found to produce a 2-V peak-to-peak 120-Hz hum. We wish to reduce the output hum to less than 1 mV without changing the signal gain. To this end, we precede the power stage with a preamplifier stage with gain a_1 and then apply negative feedback around the composite amplifier. What are the required values of a_1 and b?

1.53 (*a*) Develop an expression of the type of Eq. (1.52) for the shunt-series circuit of Fig. 1.30*b*. Then, extract *exact* expressions for a_ε and b, and show that for $a_\varepsilon \gg 1$ the expressions reduce to those of Eq. (1.62). (*b*) Repeat the calculations of Example 1.9, but using the exact expressions, and compare with the results of the example.

1.54 In the series-series circuit of Fig. P1.54 let $a = 10^4$ V/V, $R_1 = 1$ kΩ, $R_2 = 2$ kΩ, and $R_3 = 3$ kΩ. Use direct analysis to obtain an expression of the type $i_O = A_g v_I - v_L/R_o$. What are the values of A_g and R_o?

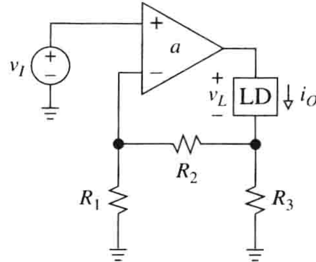


FIGURE P1.54

1.55 In the analysis leading to Eq. (1.65) we deliberately omitted r_d for simplicity. If we include r_d, then Eq. (1.65) changes to $R_i = r_d//[(R + r_o)/(1 + a)]$. Alternatively, we can adapt Eq. (1.69) and write $R_i = r_0/(1 + L)$. Find r_0 and L, and prove that the two methods provide the same result, as it should be.

1.56 Using the op amp model of Fig. 1.4 with $a = 10^4$ V/V, find the resistance seen by the source v_I and that seen by the source i_I in Fig. 1.29 if $R_1 = 1.0$ kΩ and $R_2 = 100$ kΩ. Why the difference?

1.57 In the series-series circuit of Fig. 1.30*a* let $a = 10^3$ V/V and $R = 1.0$ kΩ. (*a*) Assuming the load is a voltage source v_L (positive at the top), calculate a_ε, b, L, A_g and the resistance R_o seen by the load. (*b*) Assuming $v_I = 1.0$ V, find i_O for $v_L = 0$, $v_L = 5.0$ V, and $v_L = -4.0$ V.

1.58 Use direct analysis on the noninverting amplifier of Fig. 1.35*a* to show that its input and output resistances admit the exact expressions

$$R_i = r_d\left[1 + \frac{a}{1 + (R_2 + r_o)/R_1}\right] + [R_1//(R_2 + r_o)]$$

$$R_o = \frac{r_o}{1 + [a + r_o/(R_1//r_d)]/[1 + R_2/(R_1//r_d)]}$$

How do these expressions simplify for a well-designed inverting amplifier?

1.59 Use direct analysis on the inverting amplifier of Fig. 1.36*a* to show that its input and output resistances admit the exact expressions

$$R_i = R_1 + \frac{R_2 + r_o}{1 + a + (R_2 + r_o)/r_d} \qquad R_o = \frac{r_o}{1 + [a + r_o/(R_1//r_d)]/[1 + R_2/(R_1//r_d)]}$$

How do these expressions simplify for a well-designed inverting amplifier?

1.60 Let a voltage follower be implemented with an op amp having $r_d = 1$ kΩ, $r_o = 20$ kΩ, and $a = 10^6$ V/V (poor resistances, but excellent gain). Find A, R_i, and R_o, and comment on your findings.

1.7 The return ratio and Blackman's formula

1.61 Assuming op amps with $r_d \cong \infty$, $a = 10^3$ V/V, and $r_o \cong 0$, find the return ratio T for each of the circuits of Figs. P1.15 through P1.19.

1.62 Repeat Problem 1.61 if each source is replaced by a 10-kΩ resistor.

1.63 Let the difference amplifier of Fig. 1.18 be implemented with four matched 10-kΩ resistances and an op amp having $r_d \cong \infty$, $a = 10^2$ V/V, and $r_o = 100$ Ω. (*a*) Writing $v_O = A_2 v_2 - A_1 v_1$, use return-ratio analysis to find A_1 and $A_2 (A_1 \neq A_2$ because of feedthrough). (*b*) A figure of merit of a difference amplifier is the common-mode rejection ratio, which in this case we define as CMRR $= 20 \log |A/\Delta A|$, where $A = (A_1 + A_2)/2$ and $\Delta A = A_1 - A_2$. What is the CMRR of this circuit? What happens if a is raised to 10^3 V/V?

1.64 (*a*) Assuming the op amp of Fig. P1.64 has $r_d \cong \infty$, $a = 10^3$ V/V, and $r_o \cong 0$, and all resistances are identical, find A_{ideal} as well as the gain error GE. (*b*) Find $a_{\min}$ for GE ≤ 0.1 %.

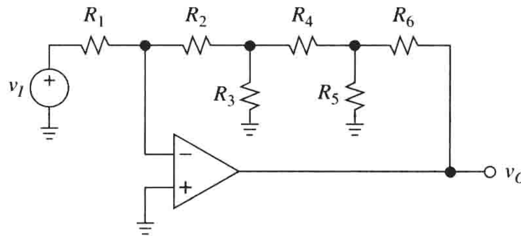


FIGURE P1.64

1.65 Let the op amp of Fig. P1.65 have $r_d \cong \infty$, $a = 10^4$ V/V, and $r_o \cong 0$. Assuming $R_1 = R_3 = R_5 = 10$ kΩ and $R_2 = R_4 = 20$ kΩ, use return-ratio analysis to find the gain $A_r = v_O/i_I$ and the resistance R_i seen by the i_I source.

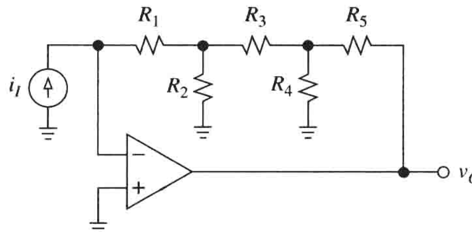


FIGURE P1.65

1.66 In the circuit[8] of Fig. P1.66 let $r_d = 50$ kΩ, $g_m = 1$ mA/V, and $r_o = 1$ MΩ. Use return-ratio analysis to find the closed-loop gain $A_r = v_o/i_s$ if $R_s = 200$ kΩ and $R_F = 100$ kΩ. Verify with PSpice. What is the percentage contribution to A_r due to feedthrough? *Hint:* Suppress the sources, apply a test current i_t out of the output node, and obtain the return ratio as $-g_m v_d / i_t$.

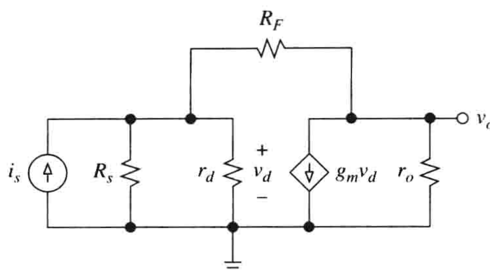


FIGURE P1.66

1.67 With reference to the circuit of Fig. P1.66, use Blackman's formula to find the resistance R_i between the input node and ground, and the resistance R_o between the output node and ground, and verify with PSpice. Assume the data of Problem 1.66, and use the hint provided there.

1.68 In the circuit of Fig. 1.42*a* let X denote the node shared by R_2, R_3, and R_4. (*a*) Assuming the data of Example 1.14, use Blackman's formula to find the resistance R_X between node X and ground. Is it large or small? Justify intuitively. (*b*) Use intuitive reasoning (no formulas) to estimate the resistance R_{XN} between node X and the inverting-input node, and the resistance R_{XO} between node X and the output node. *Hint:* What happens if you apply a test voltage between the nodes under consideration?

1.69 Redraw the circuit of Fig. P1.64, but with R_6 replaced by a load box LD as in Fig. P1.54, and let the output be the load current i_O flowing from left to right. Assuming $r_d \cong \infty$, $a = 5{,}000$ V/V, and $r_o \cong 0$, along with $R_1 = R_2 = 200\ \text{k}\Omega$, $R_3 = 100\ \text{k}\Omega$, $R_4 = 120\ \text{k}\Omega$, and $R_5 = 1.0\ \text{k}\Omega$, use return-ratio analysis to estimate the gain $A_g = i_O/v_I$ if the load LD is a short circuit. What is the resistance R_o seen by the load?

1.70 In the circuit of Fig. P1.51*a* let $a_1 = 3000$ V/V and $R_L = 2\ \text{k}\Omega$, and suppose an additional 10-kΩ resistor is connected from node v_1 to node v_O. (*a*) Sketch and label the open-loop VTC of the overall circuit, that is, the plot of v_O versus the input difference $v_D = v_P - v_N$. (*b*) Sketch and label the loop gain T versus v_I over the range $-0.3\ \text{V} \le v_I \le 0.3$ V. (*c*) Sketch and label, versus time, v_I, v_O, v_1, and v_D if v_I is a triangular wave with ±0.3-V peak values.

1.8 Op amp powering

1.71 Repeat Example 1.15, but with $v_I = -5$ V.

1.72 Assuming $I_Q = 1.5$ mA in the circuit of Fig. P1.72, calculate all currents and voltages, as well as the power dissipated inside the op amp, if (*a*) $v_I = +2$ V; (*b*) $v_I = -2$ V.

1.73 (*a*) Assuming ±15-V power supplies, design a variable voltage source over the range $0\ \text{V} \le v_S \le 10$ V. (*b*) Assuming a 1-kΩ grounded load and $I_Q = 1.5$ mA, find the maximum internal power dissipation of your op amp.

1.74 (*a*) Assuming $I_Q = 50\ \mu\text{A}$ and a grounded load of 100 kΩ at the output of the dc-offsetting amplifier of Fig. 1.17, find the values of v_I for which the op amp dissipates the maximum power. Show all corresponding voltages and currents. (*b*) Assuming

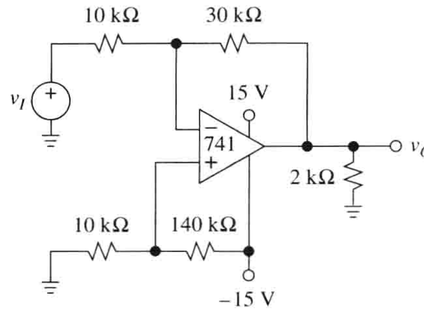


FIGURE P1.72

$\pm V_{sat} = \pm 12$ V, find the range of values of v_I for which the op amp still operates within the linear region.

1.75 In the amplifier of Fig. 1.18 let $R_1 = 30$ kΩ, $R_2 = 120$ kΩ, $R_3 = 20$ kΩ, and $R_4 = 30$ kΩ, and let the op amp be a 741-type powered from ±15 V. (*a*) If $v_2 = 2 \sin \omega t$ V, find the range of values of v_1 for which the amplifier still operates in the linear region. (*b*) If $v_1 = V_m \sin \omega t$ and $v_2 = -1$ V, find the maximum value of V_m for which the op amp still operates in the linear region. (*c*) Repeat (*a*) and (*b*) for the case in which the power supplies are lowered to ±12 V.

1.76 Redraw the circuit of Fig. P1.16, but with the source i_S replaced by a voltage source v_S, positive at the top. (*a*) How is v_O related to v_S when the op amp operates in the linear region? Assuming the op amp saturates at ±10 V, find v_N, v_P, and v_O if (*b*) $v_S = 5$ V, and (*c*) $v_S = 15$ V.

1.77 Assuming the op amp of Fig. P1.17 saturates at ±10 V, sketch and label v_N, v_P, and v_O versus time if v_S is a sinusoidal wave of ±9-V peak values.

1.78 The noninverting amplifier of Fig. 1.15*a* is implemented with $R_1 = 10$ kΩ and $R_2 = 15$ kΩ, and a 741 op amp powered from ±12-V supplies. If the circuit includes also a third 30-kΩ resistor connected between the inverting input and the 12-V supply, find v_O and v_N if (*a*) $v_I = 4$ V, and (*b*) $v_I = -2$ V.

1.79 Suppose the circuit of Fig. 1.42*a* uses 10-kΩ resistances throughout, and the op amp has $a = 10^4$ V/V and saturates at ±5 V. Assuming a sinusoidal input $v_I = V_{im} \sin(\omega t)$, and letting v_X denote the voltage at the node common to R_2, R_3, and R_4, sketch and label v_I, v_N, v_O, and v_X versus time for the following cases: (*a*) $V_{im} = 1.0$ V and (*b*) $V_{im} = 2.0$ V.

1.80 Suppose the integrator of Fig. 1.20 uses $R = 30$ kΩ, $C = 20$ nF, and an op amp that saturates at ±5 V. (*a*) Assuming C is initially discharged, find the time at which the op amp reaches saturation if v_I is changed from 0 V to 3 V at $t = 0$ (use the hint of Problem 1.42). (*b*) Sketch and label v_N both before and after saturation, and comment.

1.81 The circuit of Fig. P1.81, called a *bridge amplifier*, allows one to double the linear output range as compared with a single op amp. (*a*) Show that if the resistances are in the ratios shown, then $v_O/v_I = 2A$. (*b*) If the individual op amps saturate at ±13 V, what is the maximum peak-to-peak output voltage that the circuit can provide without distortion?

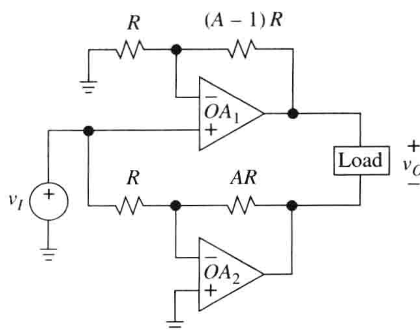


FIGURE P1.81

1.82 For the circuit of Fig. P1.72 sketch and label v_I, v_N, and v_O versus time if v_I is a triangular wave with ±5-V peak values.

REFERENCES

1. W. Jung, *Op Amp Applications Handbook* (Analog Devices Series), Elsevier/Newnes, New York, 2005, ISBN 0-7506-7844-5.
2. P. R. Gray, P. J. Hurst, S. H. Lewis, and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 5th ed., John Wiley & Sons, New York, 2009, ISBN 978-0-470-24599-6.
3. S. Franco, *Analog Circuit Design—Discrete and Integrated,* McGraw-Hill, New York, 2014.
4. H. W. Bode, *Network Analysis and Feedback Amplifier Design*, Van Nostrand, New York, 1945.
5. R. D. Middlebrook, "The General Feedback Theorem: A Final Solution for Feedback Systems," *IEEE Microwave Magazine*, April 2006, pp. 50–63.
6. R. B. Blackman, "Effect of Feedback on Impedance," *Bell Sys. Tech. J.*, Vol. 23, October 1943, pp. 269–277.
7. S. Rosenstark, *Feedback Amplifier Principles*, MacMillan, New York, 1986, ISBN 978-0672225451.
8. P. J. Hurst, "A Comparison of Two Approaches to Feedback Circuit Analysis," *IEEE Trans. on Education*, Vol. 35, No. 3, August 1992, pp. 253–261.

APPENDIX 1A STANDARD RESISTANCE VALUES

As a good work habit, always specify standard resistance values for the circuits you design (see Table 1A.1). In many applications 5% resistors are adequate; however, when higher precision is required, 1% resistors should be used. When even this tolerance is insufficient, the alternatives are either 0.1% (or better) resistors, or less precise resistors in conjunction with variable ones (trim pots) to allow for exact adjustments.

The numbers in the table are multipliers. For instance, if the calculations yield a resistance of 3.1415 kΩ, the closest 5% value is 3.0 kΩ and the closest 1% value

is 3.16 kΩ. In the design of low-power circuits, the best resistance range is usually between 1 kΩ and 1 MΩ. Try to avoid excessively high resistances (e.g., above 10 MΩ), because the stray resistance of the surrounding medium tends to decrease the effective value of your resistance, particularly in the presence of moisture and salinity. Low resistances, on the other hand, cause unnecessarily high-power dissipation.

TABLE 1A.1
Standard resistance values

5% resistor values	1% resistor values			
10	100	178	316	562
11	102	182	324	576
12	105	187	332	590
13	107	191	340	604
15	110	196	348	619
16	113	200	357	634
18	115	205	365	649
20	118	210	374	665
22	121	215	383	681
24	124	221	392	698
27	127	226	402	715
30	130	232	412	732
33	133	237	422	750
36	137	243	432	768
39	140	249	442	787
43	143	255	453	806
47	147	261	464	825
51	150	267	475	845
56	154	274	487	866
62	158	280	499	887
68	162	287	511	909
75	165	294	523	931
82	169	301	536	953
91	174	309	549	976

2

CIRCUITS WITH RESISTIVE FEEDBACK

本章将研究运算放大器电路，重点介绍它的实际应用。电路被设计成具有线性传输特性，并与频率无关。刻意设计成与频率相关的线性电路是滤波器，这些电路将在第 3 章和第 4 章进行研究。此外，非线性运算放大器电路将在第 9 章和第 13 章进行研究。

要了解一个给定电路的工作过程，首先要用理想运算放大器模型去分析，然后利用 1.6 节和 1.7 节的内容，进一步研究运算放大器的非理想特性，尤其是有限的开环增益如何影响它的闭环参数。在充分掌握了简单运算放大器模型电路的分析后，第 5 章和第 6 章将更系统地研究运算放大器的非理想特性，例如静态和动态误差等问题。在本章和其他章节中最直接受这些限制因素影响的电路将在以后进行详细分析。

本章重点

本章的前半部分再次提到 1.6 节中的 4 种反馈电路结构，来研究一些实际应用。尽管本质上一个电压型放大器也能构成跨阻放大器或 *I-V* 转换器，以及跨导放大器或 *V-I* 转换器和电流放大器。这样卓越的多样性源于负反馈能力，它能调节闭环电阻和稳定性增益。巧妙地利用这一特性使我们能理解表 1.1 所示的理想放大器的工作状态。

本章后半部分着重介绍仪表的概念和应用。研究的电路包括差分放大器、仪表放大器和传感器桥式放大器，这些电路广泛应用于当今的自动化测试、检测和控制仪器等领域。

2.1 CURRENT-TO-VOLTAGE CONVERTERS

A *current-to-voltage converter* (*I-V* converter), also called a *transresistance amplifier,* accepts an input current i_I and yields an output voltage of the type $v_O = Ai_I$, where A is the gain of the circuit in volts per ampere. Referring to Fig. 2.1, assume first that the op amp is ideal. Summing currents at the virtual-ground node gives $i_I + (v_O - 0)/R = 0$, or

$$v_O = -Ri_I \tag{2.1}$$

The gain is $-R$ and is negative because of the choice of the reference direction of i_I; inverting this direction gives $v_O = Ri_I$. The magnitude of the gain is also called the *sensitivity* of the converter because it gives the amount of output voltage change for a given input current change. For instance, for a sensitivity of 1 V/mA we need $R = 1$ kΩ, for a sensitivity of 1 V/μA we need $R = 1$ MΩ, and so on. If desired, gain can be made variable by implementing R with a potentiometer. Note that the feedback element need not necessarily be limited to a resistance. In the more general case in which it is an impedance $Z(s)$, where s is the complex frequency, Eq. (2.1) takes on the Laplace-transform form $V_o(s) = -Z(s)I_i(s)$, and the circuit is called a *transimpedance amplifier.*

We observe that the op amp eliminates loading both at the input and at the output. In fact, should the input source exhibit some finite parallel resistance R_s, the

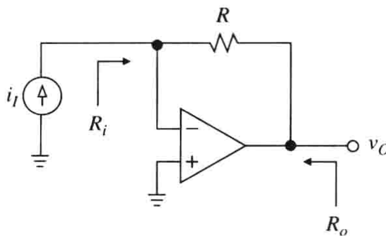


FIGURE 2.1
Basic *I-V* converter.

op amp eliminates any current loss through it by forcing 0 V across it. Also, the op amp delivers v_O to an output load R_L with zero output resistance.

Closed-Loop Parameters

Let us now investigate the departure from ideal if a practical op amp is used. Comparing with Fig. 1.28*b*, we recognize the *shunt-shunt* topology. We can thus apply the techniques of Section 1.7 and write

$$T = \frac{ar_d}{r_d + R + r_o} \tag{2.2}$$

$$A = -R\frac{1}{1+1/T} \qquad R_i = \frac{r_d \parallel (R + r_o)}{1+T} \qquad R_o \cong \frac{r_o}{1+T} \tag{2.3}$$

EXAMPLE 2.1. Find the closed-loop parameters of the circuit of Fig. 2.1 if it is implemented with a 741 op amp and $R = 1\ \text{M}\Omega$.

Solution. Substituting the given component values, we get $T = 133{,}330$, $A = -0.999993\ \text{V}/\mu\text{A}$, $R_i = 5\ \Omega$, and $R_o \cong 56\ \text{m}\Omega$.

High-Sensitivity *I-V* Converters

It is apparent that high-sensitivity applications may require unrealistically large resistances. Unless proper circuit fabrication measures are adopted, the resistance of the surrounding medium, being in parallel with R, will decrease the net feedback resistance and degrade the accuracy of the circuit. Figure 2.2 shows a widely used technique to avoid this drawback. The circuit utilizes a T-network to achieve high sensitivity without requiring unrealistically large resistances.

Summing currents at node v_1 yields $-v_1/R - v_1/R_1 + (v_O - v_1)/R_2 = 0$. But $v_1 = -Ri_I$, by Eq. (2.1). Eliminating v_1 yields

$$v_O = -kRi_I \tag{2.4a}$$

$$k = 1 + \frac{R_2}{R_1} + \frac{R_2}{R} \tag{2.4b}$$

The circuit in effect increases R by the multiplicative factor k. We can thus achieve a high sensitivity by starting out with a reasonable value of R and then multiplying it by the needed amount k.

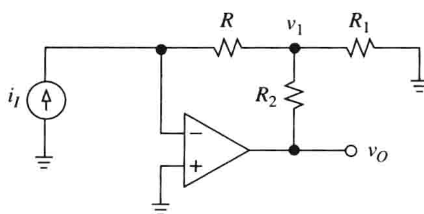


FIGURE 2.2
High-sensitivity I-V converter.

EXAMPLE 2.2. In the circuit of Fig 2.2 specify suitable component values to achieve a sensitivity of 0.1 V/nA.

Solution. We have $kR = 0.1/10^{-9} = 100\ \text{M}\Omega$, a fairly large value. Start out with $R = 1\ \text{M}\Omega$ and then multiply it by 100 to meet the specifications. Thus, $1 + R_2/R_1 + R_2/10^6 = 100$. Since we have one equation but two unknowns, fix one unknown; for example, let $R_1 = 1\ \text{k}\Omega$. Then, imposing $1 + R_2/10^3 + R_2/10^6 = 100$ yields $R_2 \cong 99\ \text{k}\Omega$ (use 100 kΩ, the closest standard). If desired, R_2 can be made variable for the exact adjustment of kR.

Real-life op amps do draw a small current at their input terminals. Called the *input bias current,* it may degrade the performance of high-sensitivity *I-V* converters, in which i_I itself is quite small. This drawback can be avoided by using op amps specifically rated for low-input bias current, such as JFET-input and MOSFET-input op amps.

Photodetector Amplifiers

One of the most frequent *I-V* converter applications is in connection with current-type photodetectors such as photodiodes and photomultipliers.[1] Another common application, *I-V* conversion of current-output digital-to-analog converters, will be discussed in Chapter 12.

Photodetectors are transducers that produce electrical current in response to incident light or other forms of radiation, such as X-rays. A transresistance amplifier is then used to convert this current to a voltage, as well as eliminate possible loading both at the input and at the output.

One of the most widely used photodetectors is the *silicon photodiode.* The reasons for its popularity are its solid-state reliability, low cost, small size, and low power dissipation.[1] The device can be used either with a reverse bias voltage, in the *photoconductive* mode, shown in Fig. 2.3*a*, or with zero bias, in the *photovoltaic* mode, shown in Fig. 2.3*b*. The photoconductive mode offers higher speed; it is therefore better suited to the detection of high-speed light pulses and to high-frequency light-beam modulation applications. The photovoltaic mode offers lower noise and is therefore better suited to measurement and instrumentation applications. The circuit of Fig. 2.3*b* can be used as a light meter by calibrating its output directly in units of light intensity.

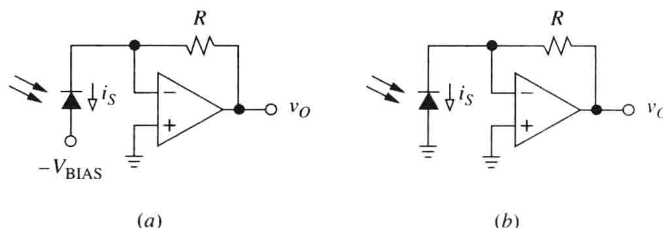


FIGURE 2.3
(*a*) Photoconductive and (*b*) photovoltaic detectors.

2.2 VOLTAGE-TO-CURRENT CONVERTERS

A *voltage-to-current converter* (*V-I* converter), also called a *transconductance amplifier,* accepts an input voltage v_I and yields an output current of the type $i_O = Av_I$, where A is the *gain,* or *sensitivity,* of the circuit, in amperes per volt. For a practical converter, the characteristic takes on the more realistic form

$$i_O = Av_I - \frac{1}{R_o}v_L \tag{2.5a}$$

where v_L is the voltage developed by the output load in response to i_O, and R_o is the converter's output resistance as seen by the load. For true *V-I* conversion, i_O must be independent of v_L; that is, we must have

$$R_o = \infty \tag{2.5b}$$

Since it outputs a current, the circuit needs a load in order to work; leaving the output port open would result in circuit malfunction as i_O would have no path in which to flow. The *voltage compliance* is the range of permissible values of v_L for which the circuit still works properly, before the onset of any saturation effects on the part of the op amp.

If both terminals of the load are uncommitted, the load is said to be of the *floating* type. Frequently, however, one of the terminals is already committed to ground or to another potential. The load is then said to be of the *grounded* type, and the current from the converter must be fed to the uncommitted terminal.

Floating-Load Converters

Figure 2.4 shows two basic implementations, both of which use the load itself as the feedback element; if one of the load terminals were already committed, it would of course not be possible to use the load as the feedback element.

In the circuit of Fig. 2.4*a* the op amp outputs whatever current i_O it takes to make the inverting-input voltage follow v_I, or to make $Ri_O = v_I$. Solving for i_O yields

$$i_O = \frac{1}{R}v_I \tag{2.6}$$

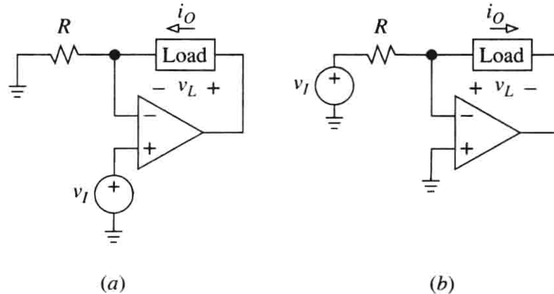


FIGURE 2.4
Floating-load *V-I* converters.

This expression holds regardless of the type of load: it can be linear, as for a resistive transducer; it can be nonlinear, as for a diode; it can have time-dependent characteristics, as for a capacitor. No matter what the load, the op amp will force it to carry the current of Eq. (2.6), which depends on the control voltage v_I and the current-setting resistance R, but not on the load voltage v_L. To achieve this goal, the op amp must swing its output to the value $v_O = v_I + v_L$, something it will readily do as long as $V_{OL} < v_O < V_{OH}$. Consequently, the voltage compliance of the circuit is $(V_{OL} - v_I) < v_L < (V_{OH} - v_I)$.

In the circuit of Fig. 2.4*b* the op amp keeps its inverting input at 0 V. Consequently, its output terminal must draw the current $i_O = (v_I - 0)/R$, and it must swing to the voltage $v_O = -v_L$. Apart from the polarity reversal, the current is the same as in Eq. (2.6); however, the voltage compliance is now $V_{OL} < v_L < V_{OH}$.

We observe that Eq. (2.6) holds for both circuits regardless of the polarity of v_I. The arrows of Fig. 2.4 show current direction for $v_I > 0$; making $v_I < 0$ will simply reverse the direction. The two converters are thus said to be *bidirectional.*

Of special importance is the case in which the load is a capacitor, so that the circuit is the familiar integrator. If v_I is kept constant, the circuit will force a constant current through the capacitor, causing it to charge or discharge, depending on the polarity of v_I, at a constant rate. This forms the basis of waveform generators such as sawtooth and triangular waveform generators, *V-F* and *F-V* converters, and dual-ramp A-D converters.

A drawback of the converter of Fig. 2.4*b* is that i_O must come from the source v_I itself, whereas in Fig. 2.4*a* the source sees a virtually infinite input resistance. This advantage, however, is offset by a more restricted voltage compliance. The maximum current either circuit can deliver to the load depends on the op amp. For the 741, this is typically 25 mA. If larger currents are required, one can either use a power op amp or a low-power op amp with an output current booster.

EXAMPLE 2.3. Let both circuits of Fig. 2.4 have $v_I = 5$ V, $R = 10$ kΩ, $\pm V_{sat} = \pm 13$ V, and a resistive load R_L. For both circuits find (*a*) i_O; (*b*) the voltage compliance; (*c*) the maximum permissible value of R_L.

Solution.

(*a*) $i_O = 5/10 = 0.5$ mA, flowing from right to left in the circuit of Fig. 2.4*a* and from left to right in that of Fig. 2.4*b*.
(*b*) For the circuit of Fig. 2.4*a*, $-18\text{ V} < v_L < 8\text{ V}$; for the circuit of Fig. 2.4*b*, $-13\text{ V} < v_L < 13\text{ V}$.
(*c*) With a purely resistive load, v_L will always be positive. For the circuit of Fig. 2.4*a*, $R_L < 8/0.5 = 16$ kΩ; for the circuit of Fig. 2.4*b*, $R_L < 13/0.5 = 26$ kΩ.

Practical Op Amp Limitations

We now wish to investigate the effect of using a practical op amp. After the op amp is replaced with its practical model, the circuit of Fig. 2.4*a* becomes as in Fig. 2.5. Summing voltages, we get $v_I - v_D + v_L + r_o i_O - a v_D = 0$. Summing currents, $i_O + v_D/r_d - (v_I - v_D)/R = 0$. Eliminating v_D and rearranging, we can put i_O

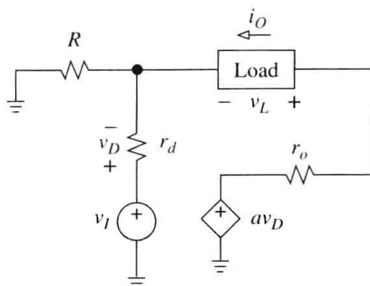


FIGURE 2.5
Investigating the effect of using a practical op amp.

in the form of Eq. (2.5*a*) with

$$A = \frac{1}{R}\frac{a - R/r_d}{1 + a + r_o/R + r_o/r_d} \qquad R_o = (R \parallel r_d)(1 + a) + r_o \qquad (2.7)$$

It is apparent that as $a \to \infty$, we get the ideal results $A \to 1/R$ and $R_o \to \infty$. However, for a finite gain a, A will exhibit some error, and R_o, though large, will not be infinite, indicating a weak dependence of i_O on v_L. Similar considerations hold for the circuit of Fig. 2.4*b*.

Grounded-Load Converters

When one of its terminals is already committed, the load can no longer be placed within the feedback loop of the op amp. Figure 2.6*a* shows a converter suitable for grounded loads. Referred to as the *Howland current pump* after its inventor, the circuit consists of an input source v_I with series resistance R_1, and a negative-resistance converter synthesizing a grounded resistance of value $-R_2R_3/R_4$. The circuit seen by the load admits the Norton equivalent of Fig. 2.6*b*, whose *i*-*v* characteristic is given by Eq. (2.5*a*). We wish to find the overall output resistance R_o seen by the load.

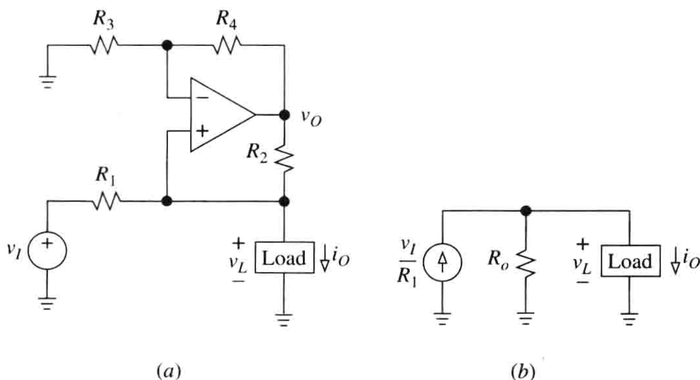


FIGURE 2.6
Howland current pump and its Norton equivalent.

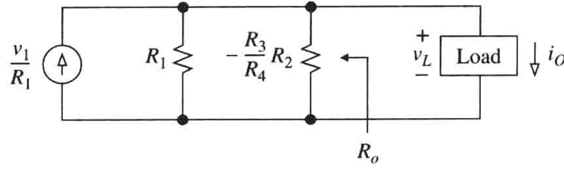


FIGURE 2.7
Using a negative resistance to control R_o.

To this end, we first perform a source transformation on the input source v_I and its resistance R_1, and then we connect the negative resistance in parallel, as depicted in Fig. 2.7. We have $1/R_o = 1/R_1 + 1/(-R_2R_3/R_4)$. Expanding and rearranging, we get

$$R_o = \frac{R_2}{R_2/R_1 - R_4/R_3} \tag{2.8}$$

As we know, for true current-source behavior we must have $R_o = \infty$. To achieve this condition, the four resistances must form a *balanced bridge:*

$$\frac{R_4}{R_3} = \frac{R_2}{R_1} \tag{2.9}$$

When this condition is met, the output becomes independent of v_L:

$$i_O = \frac{1}{R_1}v_I \tag{2.10}$$

Clearly, the gain of the converter is $1/R_1$. For $v_I > 0$ the circuit will *source* current to the load, and for $v_I < 0$ it will *sink* current. Since $v_L = v_O R_3/(R_3 + R_4) = v_O R_1/(R_1 + R_2)$, the voltage compliance is, assuming symmetric output saturation,

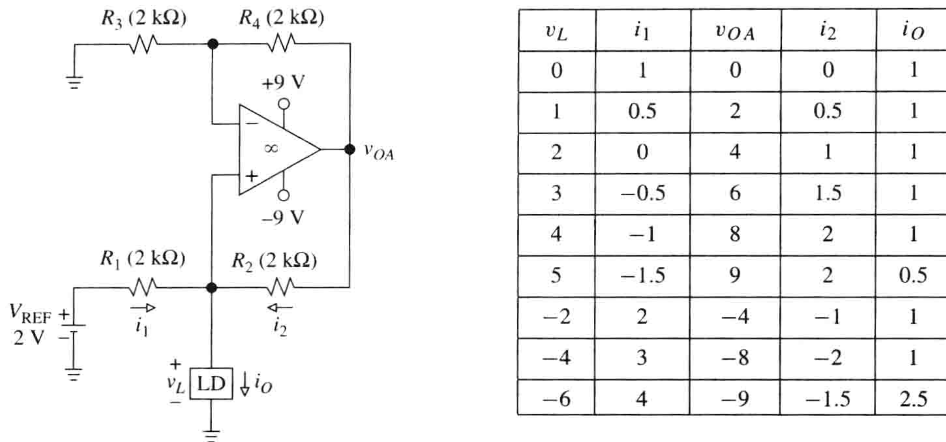
$$|v_L| \leq \frac{R_1}{R_1 + R_2}V_{\text{sat}} \tag{2.11}$$

For the purpose of extending the compliance, it is thus desirable to keep R_2 sufficiently smaller than R_1 (e.g., $R_2 \cong 0.1R_1$).

EXAMPLE 2.4. The Howland pump of Fig. 2.8 uses a 2-V voltage reference to source a stable current of 1.0 mA. Assuming a rail-to-rail op amp ($\pm V_{\text{sat}} = \pm 9$ V), prepare a table showing all voltages and currents for $v_L = 0, 1, 2, 3, 4, 5, -2, -4$, and -6 V, and give a word description of circuit behavior. What is the voltage compliance of this pump?

Solution. So long as $-9 \text{ V} \leq v_{OA} \leq +9$ V, the op amp will operate in the linear region to give $v_{OA} = 2v_L$. The voltage compliance is thus $-4.5 \text{ V} \leq v_L \leq +4.5$ V. By KCL we have $i_O = i_1 + i_2$, where

$$i_1 = \frac{V_{\text{REF}} - v_L}{R_1} \qquad i_2 = \frac{v_{OA} - v_L}{R_2} = \frac{2v_L - v_L}{R_2} = \frac{v_L}{R_2}$$



v_L	i_1	v_{OA}	i_2	i_O
0	1	0	0	1
1	0.5	2	0.5	1
2	0	4	1	1
3	−0.5	6	1.5	1
4	−1	8	2	1
5	−1.5	9	2	0.5
−2	2	−4	−1	1
−4	3	−8	−2	1
−6	4	−9	−1.5	2.5

FIGURE 2.8
Current source of Example 2.4, and distribution of voltages/currents for different load voltages (voltages are in V, currents in mA).

so, for linear-region operation, $i_1 + i_2 = V_{REF}/R_1$. Plugging in the given v_L values we readily fill in the first five data rows of the table of Fig. 2.8. Thus, for $v_L = 0$, i_O comes entirely from V_{REF}, but as the voltage v_L developed by the load increases, the contribution by V_{REF} decreases while that by v_{OA} increases in such a way as to add up to $i_O = 1$ mA regardless of v_L. (Note that for $v_L > V_{REF}$, i_1 actually changes polarity!) However, for $v_L > 9/2 = 4.5$ V, the op amp saturates, ceasing to provide the required regulation (indeed, for $v_L = 5$ V, i_O drops to 0.5 mA).

For $v_L < 0$, v_{OA} becomes negative, pulling current out of the load node ($i_2 < 0$) so as to compensate for the fact that now $i_1 > 1$ mA. For $v_L < -9/2 = -4.5$ V, the op amp again saturates and stops regulating. It is fascinating how the op amp tries to provide whatever voltage and current it takes to ensure $i_O = 1$ mA regardless of v_L (this, of course, so long as it manages to stay out of saturation).

We observe that the Howland pump comprises both a *negative* and a *positive* feedback path. Denoting the load resistance as R_L, we find the loop gain via Eq. (1.78) as

$$T = \frac{a(v_N - v_P)}{v_T} = a\left(\frac{R_3}{R_3 + R_4} - \frac{R_1//R_L}{R_1//R_L + R_2}\right)$$
$$= a\left(\frac{1}{1 + R_2/R_1} - \frac{1}{1 + R_2/R_1 + R_2/R_L}\right)$$

where Eq. (2.9) has been used. It is apparent that as long as the circuit is terminated on some finite load $0 \leq R_L < \infty$, we have $T > 0$, indicating that negative feedback *prevails* over positive feedback, thus ensuring a stable circuit.

Effect of Resistance Mismatches

In a practical circuit the resistive bridge is likely to be unbalanced because of resistance tolerances. This will inevitably degrade R_o, which should be infinite for true

current-source behavior. It is therefore of interest to estimate the worst-case value of R_o for given resistance-tolerance specifications.

An unbalanced bridge implies unequal resistance ratios in Eq. (2.9), a condition that we can express in terms of the *imbalance factor* ϵ as

$$\frac{R_4}{R_3} = \frac{R_2}{R_1}(1-\epsilon) \tag{2.12}$$

Substituting in Eq. (2.8) and simplifying, we obtain

$$R_o = \frac{R_1}{\epsilon} \tag{2.13}$$

As expected, the smaller the imbalance, the larger R_o. In the limit of perfect balance, or as $\epsilon \to 0$, we would of course have $R_o \to \infty$. We observe that ϵ and therefore R_o can be either positive or negative, depending on the direction in which the bridge is unbalanced. By Eq. (2.5*a*), $-1/R_o$ represents the slope of the i_O versus v_L characteristic. Consequently, $R_o = \infty$ implies a perfectly horizontal characteristic, $R_o > 0$ implies a tilt toward the right, and $R_o < 0$ implies a tilt toward the left.

EXAMPLE 2.5. (*a*) Discuss the implications of using 1% resistances in the circuit of Example 2.4. (*b*) Repeat for 0.1% resistances. (*c*) Find the resistance tolerance needed for $|R_o| \geq 10$ MΩ.

Solution. The worst-case bridge imbalance occurs when, for instance, the ratio R_2/R_1 is maximized and R_4/R_3 is minimized, that is, when R_2 and R_3 are maximized and R_1 and R_4 are minimized. Denoting the percentage tolerance of the resistances as p so that, for instance, for 1% resistances we have $p = 0.01$, we observe that to achieve the balanced condition of Eq. (2.9), the minimized resistances must be multiplied by $1+p$, and the maximized ones by $1-p$, thus giving

$$\frac{R_4(1+p)}{R_3(1-p)} = \frac{R_2(1-p)}{R_1(1+p)}$$

Rearranging, we get

$$\frac{R_4}{R_3} = \frac{R_2(1-p)^2}{R_1(1+p)^2} \cong \frac{R_2}{R_1}(1-p)^2(1-p)^2 \cong \frac{R_2}{R_1}(1-4p)$$

where we have exploited the fact that for $p \ll 1$ we can approximate $1/(1+p) \cong 1-p$ and we can ignore terms in p^n, $n \geq 2$. Comparison with Eq. (2.12) indicates that we can write

$$|\epsilon|_{\max} \cong 4p$$

(*a*) For 1% resistances we have $|\epsilon|_{\max} \cong 4 \times 0.01 = 0.04$, indicating a resistance ratio mismatch as large as 4%. Thus, $|R_o|_{\min} = R_1/|\epsilon|_{\max} \cong 2/0.04 = 50$ kΩ, indicating that with 1% resistances we can expect R_o to be anywhere in the range $|R_o| \geq 50$ kΩ.

(*b*) Improving the tolerance by an order of magnitude increases $|R_o|_{\min}$ by the same amount, so $|R_o| \geq 500$ kΩ.

(*c*) Since 0.1% resistances give $|R_o|_{\min} = 0.5$ MΩ, it follows that for $|R_o|_{\min} = 10$ MΩ we need to improve the tolerance by a factor of $10/0.5 = 20$. Consequently, $p = 0.1/20 = 0.005\%$, implying highly precise resistors!

An alternative to highly precise resistors is to make provision for resistance trimming. However, a good designer will strive to avoid trimmers whenever possible because they are mechanically and thermally unstable, they have finite resolution, and

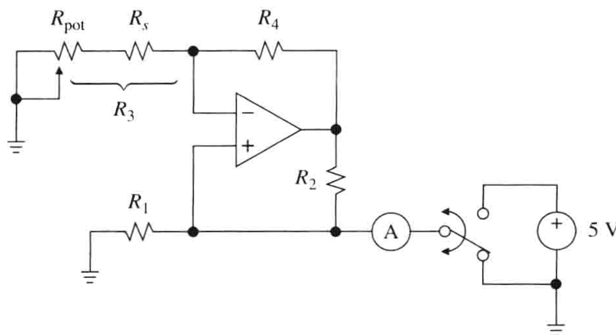


FIGURE 2.9
Howland circuit calibration.

they are bulkier than ordinary resistors. Moreover, the calibration procedure increases production costs. There are, nonetheless, situations in which, after a careful analysis of cost, complexity, and other pertinent factors, trimming still proves preferable.

Figure 2.9 shows a setup for the calibration of the Howland circuit. The input is grounded, and the load is replaced by a sensitive ammeter initially connected to ground. In this state the ammeter reading should be zero; however, because of op amp nonidealities such as the input bias current and the input offset voltage, to be discussed in Chapter 5, the reading will generally be nonzero, albeit small. To calibrate the circuit for $R_o = \infty$, we flip the ammeter to some other voltage, such as 5 V, and we adjust the wiper for the same ammeter reading as when the ammeter is connected to ground.

EXAMPLE 2.6. In the circuit of Example 2.4 specify a suitable trimmer/resistor replacement for R_3 to allow bridge balancing in the case of 1% resistances.

Solution. Since $4pR_1 = 4 \times 0.01 \times 2 \times 10^3 = 80\ \Omega$, the series resistance R_s must be less than 2.0 kΩ by at least 80 Ω. To be on the safe side, pick $R_s = 1.91$ kΩ, 1%. Then, $R_{pot} = 2(2 - 1.91)10^3 = 180\ \Omega$ (pick a 200-Ω pot).

Effect of Finite Open-Loop Gain

We now investigate the effect of a finite open-loop gain on the transfer characteristic of the Howland circuit. To evidence the effect of the op amp alone, we assume the resistances to form a perfectly balanced bridge. With reference to Fig. 2.6*a*, we have, by KCL, $i_O = (v_I - v_L)/R_1 + (v_O - v_L)/R_2$. The circuit can be viewed as a noninverting amplifier that amplifies v_L to yield $v_O = v_L a/[1 + aR_3/(R_3 + R_4)]$. Using Eq. (2.9), this can be written as $v_O = v_L a/[1 + aR_1/(R_1 + R_2)]$. Eliminating v_O and rearranging gives

$$i_O = \frac{1}{R_1}v_I - \frac{1}{R_o}v_L \qquad R_o = (R_1//R_2)\left(1 + \frac{a}{1 + R_2/R_1}\right) \tag{2.14}$$

(Note that R_o could have been derived also via Blackman's formula.) It is interesting that the effect of finite open-loop gain a is to lower R_o from ∞, but without changing the sensitivity ($A = 1/R_1$).

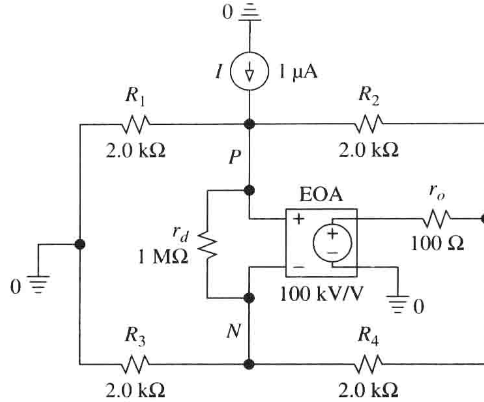


FIGURE 2.10
PSpice circuit for Example 2.7.

EXAMPLE 2.7. (*a*) Calculate R_o for the Howland pump of Example 2.4 if the op amp has $a = 10^5$ V/V, $r_d = \infty$, and $r_o = 0$, and verify with PSpice. (*b*) What happens if you include the more realistic values $r_d = 1$ MΩ and $r_o = 100$ Ω?

Solution.

(*a*) By Eq. (2.14) we have $R_o = (2//2)10^3[1 + 10^5/(1 + 2/2)] = 50$ MΩ. To verify with PSpice, apply a 1-μA test current to the P node, as shown in Fig. 2.10 (note that by PSpice convention for VCVSs, the "+" input is shown at top and the "-" input at bottom). Directing PSpice to perform the dc analysis with $r_d = \infty$ and $r_o = 0$, we get V(P) = 49.999 V, so $R_o = (49.999)/(10^{-6}) \cong 50$ MΩ, in agreement with hand calculation.

(*b*) Rerunning PSpice with $r_d = 1$ MΩ and $r_o = 100$ Ω, we get $R_o = 47.523$ MΩ, a slightly lower value because of the loop-gain reduction brought about by loading, especially at the op amp's output.

Improved Howland Current Pump

Depending on circuit conditions, the Howland circuit can be unnecessarily wasteful of power. As an example, let $v_I = 1$ V, $R_1 = R_3 = 1$ kΩ, and $R_2 = R_4 = 100$ Ω, and suppose the load is such that $v_L = 10$ V. By Eq. (2.10), $i_O = 1$ mA. Note, however, that the current through R_1 toward the left is $i_1 = (v_L - v_I)/R_1 = (10 - 1)/1 = 9$ mA, indicating that the op amp will have to waste 9 mA through R_1 to deliver only 1 mA to the load under the given conditions. This inefficient use of power can be avoided with the modification of Fig. 2.11, in which the resistance R_2 has been split into two parts, R_{2A} and R_{2B}, such that the balanced condition is now

$$\frac{R_4}{R_3} = \frac{R_{2A} + R_{2B}}{R_1} \tag{2.15a}$$

It is left as an exercise (see Problem 2.16) to prove that when this condition is met, the load still sees $R_o = \infty$, but the transfer characteristic is now

$$i_O = \frac{R_2/R_1}{R_{2B}} v_I \tag{2.15b}$$

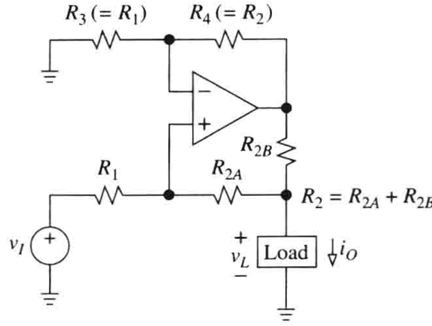


FIGURE 2.11
Improved Howland circuit.

Aside from the gain term R_2/R_1, the sensitivity is now set by R_{2B}, indicating that R_{2B} can be made as small as needed while the remaining resistances are kept high in order to conserve power. For instance, letting $R_{2B} = 1$ kΩ, $R_1 = R_3 = R_4 =$ 100 kΩ, and $R_{2A} = 100 - 1 = 99$ kΩ, we still get $i_O = 1$ mA with $v_I = 1$ V. However, even with $v_L = 10$ V, very little power is now wasted in the large 100-kΩ resistances. The voltage compliance is approximately $|v_L| \leq |V_{\text{sat}}| - R_{2B}|i_O|$. By Eq. (2.15*b*), this can be written as $|v_L| \leq |V_{\text{sat}}| - (R_2/R_1)|v_I|$.

Since Howland circuits employ both positive and negative feedback, they may become oscillatory under certain conditions. Two small capacitors (typically on the order of 10 pF) in parallel with R_4 and R_1 are usually adequate to make negative feedback prevail over positive feedback at high frequencies and thus stabilize the circuit.

2.3 CURRENT AMPLIFIERS

Even though op amps are voltage amplifiers, they can also be configured for current amplification. The transfer characteristic of a practical current amplifier is of the type

$$i_O = Ai_I - \frac{1}{R_o}v_L \qquad (2.16a)$$

where A is the gain in amperes per ampere, v_L is the output load voltage, and R_o is the output resistance as seen by the load. To make i_O independent of v_L, a current amplifier must have

$$R_o = \infty \qquad (2.16b)$$

Current-mode amplifiers are used in applications in which information is more conveniently represented in terms of current than in terms of voltage, for example, in two-wire remote sensing instrumentation, photodetector output conditioning, and V-F converter input conditioning.

Figure 2.12*a* shows a current amplifier with a floating load. Assume first that the op amp is ideal. By KCL, i_O is the sum of the currents coming from R_1 and R_2, or $i_O = i_I + (R_2 i_I)/R_1$, or $i_O = Ai_I$, where

$$A = 1 + \frac{R_2}{R_1} \qquad (2.17)$$

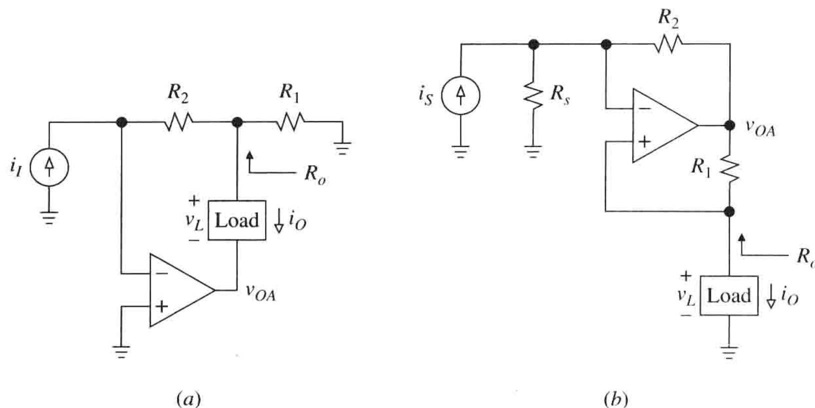


FIGURE 2.12
Current amplifiers: (*a*) floating-load type, and (*b*) grounded-load type.

This holds regardless of v_L, indicating that the circuit yields $R_o = \infty$. If the op amp has a finite gain a, one can prove (see Problem 2.24) that

$$A = 1 + \frac{R_2/R_1}{1 + 1/a} \qquad R_o = R_1(1 + a) \tag{2.18}$$

indicating a gain error as well as a finite output resistance. One can readily verify that the voltage compliance is $-(V_{OH} + R_2 i_I) \le v_L \le -(V_{OL} + R_2 i_I)$.

Figure 2.12*b* shows a grounded-load current amplifier. Because of the virtual short, the voltage across the input source is v_L, so the current entering R_2 from the left is $i_S - v_L/R_s$. By KVL, we have $v_{OA} = v_L - R_2(i_S - v_L/R_s)$. By KCL and Ohm's law, $i_O = (v_{OA} - v_L)/R_1$. Eliminating v_{OA} gives $i_O = A i_S - (1/R_o)v_L$, where

$$A = -\frac{R_2}{R_1} \qquad R_o = -\frac{R_1}{R_2} R_s \tag{2.19}$$

The negative gain indicates that the actual direction of i_O is opposite to that shown. Consequently, sourcing current to (or sinking current from) the circuit will cause it to sink current from (or source current to) the load. If $R_1 = R_2$, then $A = -1$ A/A and the circuit functions as a *current reverser,* or *current mirror.*

We observe that R_o is negative, something we could have anticipated by comparing our amplifier with the negative-resistance converter of Fig. 1.21*b*. The fact that R_o is finite indicates that i_O is not independent of v_L. To avoid this shortcoming, the circuit is used primarily in connection with loads of the virtual-ground type ($v_L = 0$), as in certain types of current-to-frequency converters and logarithmic amplifiers.

2.4 DIFFERENCE AMPLIFIERS

The difference amplifier was introduced in Section 1.4, but since it forms the basis of other important circuits, such as instrumentation and bridge amplifiers, we now wish to analyze it in greater detail. Referring to Fig. 2.13*a*, we recall that as long as

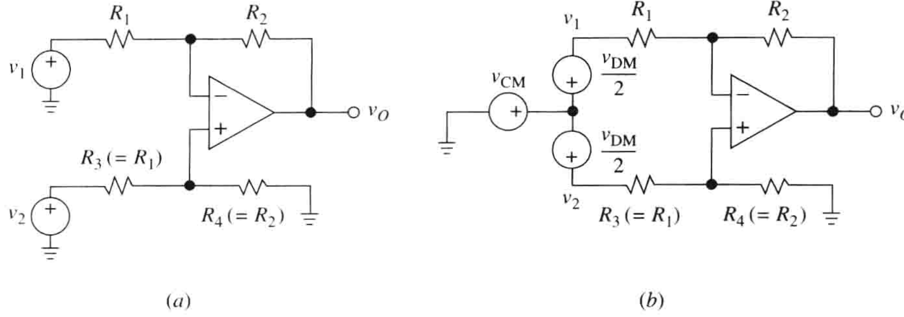


FIGURE 2.13
(*a*) Difference amplifier. (*b*) Expressing the inputs in terms of the common-mode and differential-mode components v_{CM} and v_{DM}.

the resistances satisfy the balanced-bridge condition

$$\frac{R_4}{R_3} = \frac{R_2}{R_1} \tag{2.20a}$$

the circuit is a true difference amplifier, that is, its output is linearly proportional to the difference of its inputs,

$$v_O = \frac{R_2}{R_1}(v_2 - v_1) \tag{2.20b}$$

The unique characteristics of the difference amplifier are better appreciated if we introduce the *differential-mode* and the *common-mode* input components, defined as

$$v_{DM} = v_2 - v_1 \tag{2.21a}$$

$$v_{CM} = \frac{v_1 + v_2}{2} \tag{2.21b}$$

Inverting these equations, we can express the actual inputs in terms of the newly defined components:

$$v_1 = v_{CM} - \frac{v_{DM}}{2} \tag{2.22a}$$

$$v_2 = v_{CM} + \frac{v_{DM}}{2} \tag{2.22b}$$

This allows us to redraw the circuit in the form of Fig. 2.13*b*. We can now concisely define a true difference amplifier as a circuit that responds only to the differential-mode component v_{DM}, completely ignoring the common-mode component v_{CM}. In particular, if we tie the inputs together to make $v_{DM} = 0$, and we apply a common voltage $v_{CM} \neq 0$, a true difference amplifier will yield $v_O = 0$ regardless of the magnitude and polarity of v_{CM}. Conversely, this can serve as a test for finding how close a practical difference amplifier is to ideal. The smaller the output variation is because of a given variation of v_{CM}, the closer the amplifier is to ideal.

The decomposition of v_1 and v_2 into the components v_{DM} and v_{CM} not only is a matter of mathematical convenience but also reflects a situation quite common in practice, that of a low-level differential signal riding on a high common-mode signal, as in the case of transducer signals. The useful signal is the differential one;

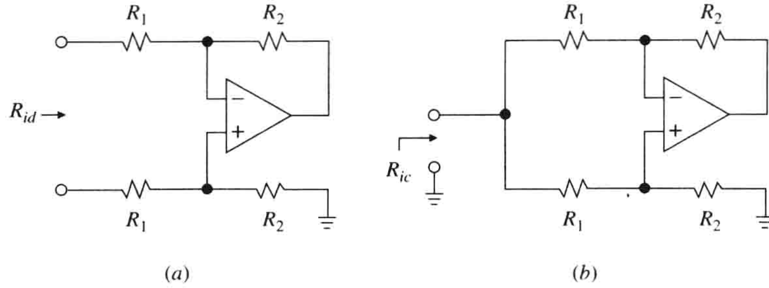


FIGURE 2.14
Differential-mode and common-mode input resistances.

extracting it from the high common-mode environment and then amplifying it can be a challenging task. Difference-type amplifiers are the natural candidates to meet this challenge.

Figure 2.14 illustrates the *differential-mode* and *common-mode input resistances*. It is readily seen (see Problem 2.30) that

$$R_{id} = 2R_1 \qquad R_{ic} = \frac{R_1 + R_2}{2} \tag{2.23}$$

Effect of Resistance Mismatches

A difference amplifier will be insensitive to v_{CM} only as long as the op amp is ideal and the resistors satisfy the balanced-bridge condition of Eq. (2.20*a*). The effect of op amp nonidealities will be investigated in Chapters 5 and 6; here we shall assume ideal op amps and explore only the effect of resistance mismatches. In general, it can be said that if the bridge is unbalanced, the circuit will respond not only to v_{DM} but also to v_{CM}.

EXAMPLE 2.8. In the circuit of Fig. 2.13*a* let $R_1 = R_3 = 10\ \text{k}\Omega$ and $R_2 = R_4 = 100\ \text{k}\Omega$. (*a*) Assuming perfectly matched resistors, find v_O for each of the following input voltage pairs: $(v_1, v_2) = (-0.1\ \text{V}, +0.1\ \text{V})$, (4.9 V, 5.1 V), (9.9 V, 10.1 V). (*b*) Repeat (*a*) with the resistors mismatched as follows: $R_1 = 10\ \text{k}\Omega$, $R_2 = 98\ \text{k}\Omega$, $R_3 = 9.9\ \text{k}\Omega$, and $R_4 = 103\ \text{k}\Omega$. Comment.

Solution.

(*a*) $v_O = (100/10)(v_2 - v_1) = 10(v_2 - v_1)$. Since $v_2 - v_1 = 0.2$ V in each of the three cases, we get $v_O = 10 \times 0.2 = 2$ V regardless of the common-mode component, which is, in order, $v_{CM} = 0$ V, 5 V, and 10 V for the three input voltage pairs.

(*b*) By the superposition principle, $v_O = A_2 v_2 - A_1 v_1$, where $A_2 = (1 + R_2/R_1)/(1 + R_3/R_4) = (1 + 98/10)/(1 + 9.9/103) = 9.853$ V/V, and $A_1 = R_2/R_1 = 98/10 = 9.8$ V/V. Thus, for $(v_1, v_2) = (-0.1\ \text{V}, +0.1\ \text{V})$ we obtain $v_O = 9.853(0.1) - 9.8(-0.1) = 1.965$ V. Likewise, for $(v_1, v_2) = (4.9\ \text{V}, 5.1\ \text{V})$ we get $v_O = 2.230$ V, and for $(v_1, v_2) = (9.9\ \text{V}, 10.1\ \text{V})$ we get $v_O = 2.495$ V. As a consequence of mismatched resistors, not only do we have $v_O \neq 2$ V, but v_O also changes with the common-mode component. Clearly the circuit is no longer a true difference amplifier.

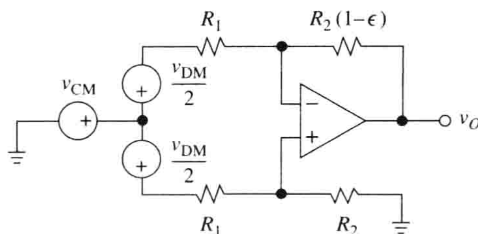


FIGURE 2.15
Investigating the effect of resistance mismatches.

The effect of bridge imbalance can be investigated more systematically by introducing the *imbalance factor* ϵ, in the manner of the Howland circuit of Section 2.2. With reference to Fig. 2.15 we conveniently assume that three of the resistances possess their nominal values while the fourth is expressed as $R_2(1-\epsilon)$ to account for the imbalance. Applying the superposition principle,

$$v_O = -\frac{R_2(1-\epsilon)}{R_1}\left(v_{CM} - \frac{v_{DM}}{2}\right) + \frac{R_1 + R_2(1-\epsilon)}{R_1} \times \frac{R_2}{R_1 + R_2}\left(v_{CM} + \frac{v_{DM}}{2}\right)$$

Multiplying out and collecting terms, we can put v_O in the insightful form

$$v_O = A_{dm} v_{DM} + A_{cm} v_{CM} \tag{2.24a}$$

$$A_{dm} = \frac{R_2}{R_1}\left(1 - \frac{R_1 + 2R_2}{R_1 + R_2}\frac{\epsilon}{2}\right) \tag{2.24b}$$

$$A_{cm} = \frac{R_2}{R_1 + R_2}\epsilon \tag{2.24c}$$

As expected, Eq. (2.24*a*) states that with an unbalanced bridge, the circuit responds not only to v_{DM} but also to v_{CM}. For obvious reasons A_{dm} and A_{cm} are called, respectively, the *differential-mode gain* and the *common-mode gain*. Only in the limit $\epsilon \to 0$ do we obtain the ideal results $A_{dm} = R_2/R_1$ and $A_{cm} = 0$.

The ratio A_{dm}/A_{cm} represents a figure of merit of the circuit and is called the *common-mode rejection ratio* (CMRR). Its value is expressed in decibels (dB) as

$$\text{CMRR}_{\text{dB}} = 20 \log_{10}\left|\frac{A_{dm}}{A_{cm}}\right| \tag{2.25}$$

For a true difference amplifier, $A_{cm} \to 0$ and thus $\text{CMRR}_{\text{dB}} \to \infty$. For a sufficiently small imbalance factor ϵ, the second term within parentheses in Eq. (2.24*b*) can be ignored in comparison with unity, and we can write $A_{dm}/A_{cm} \cong (R_2/R_1)/[R_2\epsilon/(R_1 + R_2)]$, or

$$\text{CMRR}_{\text{dB}} \cong 20 \log_{10}\left|\frac{1 + R_2/R_1}{\epsilon}\right| \tag{2.26}$$

The reason for using the absolute value is that ϵ can be positive or negative, depending on the direction of the imbalance. Note that for a given ϵ, the larger the differential gain R_2/R_1, the higher the CMRR of the circuit.

EXAMPLE 2.9. In Fig. 2.13*a* let $R_1 = R_3 = 10\ \text{k}\Omega$ and $R_2 = R_4 = 100\ \text{k}\Omega$. (*a*) Discuss the implications of using 1% resistors. (*b*) Illustrate the case in which the inputs are tied together and are driven by a common 10-V source. (*c*) Estimate the resistance tolerance needed for a guaranteed CMRR of 80 dB.

Solution.

(*a*) Proceeding along lines similar to those in Example 2.5, we can write $|\epsilon|_{\max} \cong 4p$, where p is the percentage tolerance. With $p = 1\% = 0.01$, we get $|\epsilon|_{\max} \cong 0.04$. The worst-case scenario corresponds to $A_{\text{dm(min)}} \cong (100/10)[1-(210/110)\times 0.04/2] = 9.62$ V/V $\neq 10$ V/V, and $A_{\text{cm(max)}} \cong (100/110) \times 0.04 = 0.0364 \neq 0$. Thus, $\text{CMRR}_{\min} = 20\log_{10}(9.62/0.0364) = 48.4$ dB.

(*b*) With $v_{\text{DM}} = 0$ and $v_{\text{CM}} = 10$ V, the output error can be as large as $v_O = A_{\text{cm(max)}} \times v_{\text{CM}} = 0.0364 \times 10 = 0.364\ \text{V} \neq 0$.

(*c*) To achieve a higher CMRR, we need to further decrease ϵ. By Eq. (2.26), $80 \cong 20\log_{10}[(1+10)/|\epsilon|_{\max}]$, or $|\epsilon|_{\max} = 1.1 \times 10^{-3}$. Then $p = |\epsilon|_{\max}/4 = 0.0275\%$.

It is apparent that for high CMRRs the resistors must be very tightly matched. The INA105 is a general-purpose monolithic difference amplifier[4] with four identical resistors that are matched within 0.002%. In that case, Eq. (2.26) yields $\text{CMRR}_{\text{dB}} = 100$ dB.

The CMRR of a practical amplifier can be maximized by adjusting one of its resistors, usually R_4. This is shown in Fig. 2.16. The selection of the series resistance R_s and R_{pot} follows the lines of the Howland circuit of Example 2.6. Calibration is done with the inputs tied together to eliminate v_{DM} and evidence only v_{CM}. The latter is then flipped back and forth between two predetermined values, such as -5 V and $+5$ V, and the wiper is adjusted for a minimum variation at the output. To preserve bridge balance with temperature and aging, it is advisable to use a metal-film resistor array.

So far we have assumed ideal op amps. When studying their practical limitations in Chapter 5, we shall see that op amps are themselves sensitive to v_{CM}, so the CMRR of a practical difference amplifier is actually the result of two effects: bridge imbalance and op amp nonideality. The two effects are interrelated so that it is possible to unbalance the bridge in such a way as to approximately cancel out the

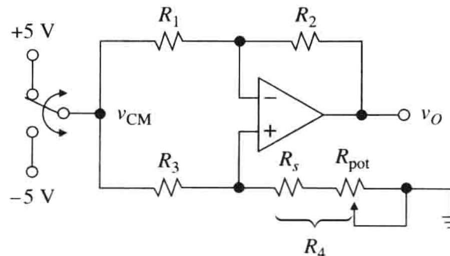


FIGURE 2.16
Difference-amplifier calibration.

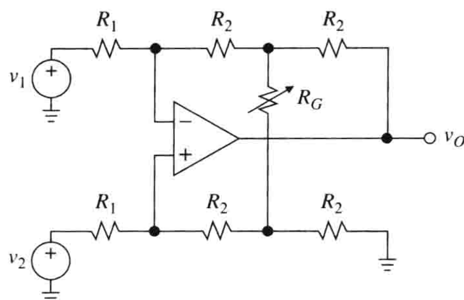


FIGURE 2.17
Difference amplifier with variable gain.

effect of the op amp. Indeed, this is what we do when we seek the minimum output variation during the calibration routine.

Variable Gain

Equation (2.20*b*) might leave the impression that gain can be varied by varying just one resistor, say, R_2. Since we must also satisfy Eq. (2.20*a*), two resistors rather than one would have to be varied, and in such a way as to maintain a very tight matching. This awkward task is avoided with the modification of Fig. 2.17, which makes it possible to vary the gain without disturbing bridge balance. It is left as an exercise (see Problem 2.31) to prove that if the various resistances are in the ratios shown, then

$$v_O = \frac{2R_2}{R_1}\left(1 + \frac{R_2}{R_G}\right)(v_2 - v_1) \tag{2.27}$$

so that gain can be varied by varying the single resistor R_G.

It is often desirable that gain vary linearly with the adjusting potentiometer to facilitate gain readings from potentiometer settings. Unfortunately, the circuit of Fig. 2.17 exhibits a nonlinear relationship between gain and R_G. This drawback is avoided by using an additional op amp, as in Fig. 2.18. As long as the closed-loop

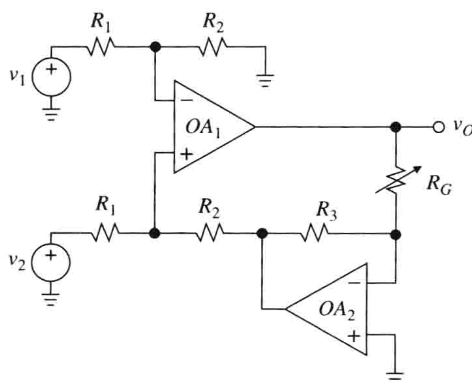


FIGURE 2.18
Difference amplifier with linear gain control.

output resistance of OA_2 is negligible, bridge balance will be unaffected. Moreover, since OA_2 provides phase reversal, the feedback signal must now be applied to the noninverting input of OA_1. One can readily prove (see Problem 2.32) that

$$v_O = \frac{R_2 R_G}{R_1 R_3}(v_2 - v_1) \tag{2.28}$$

so that the gain is linearly proportional to R_G.

Ground-Loop Interference Elimination

In practical installations source and amplifier are often far apart and share the common ground bus with a variety of other circuits. Far from being a perfect conductor, the ground bus has a small distributed resistance, inductance, and capacitance and thus behaves as a distributed impedance. Under the effect of the various currents flowing on the bus, this impedance will develop a small voltage drop, causing different points on the bus to be at slightly different potentials. In Fig. 2.19, Z_g denotes the ground-bus impedance between the input signal common N_i and the output signal common N_o, and v_g is the corresponding voltage drop. Ideally, v_g should have no effect on circuit performance.

Consider the arrangement of Fig. 2.19*a*, where v_i is to be amplified by an ordinary inverting amplifier. Unfortunately, the amplifier sees v_i and v_g in series, so

$$v_o = -\frac{R_2}{R_1}(v_i + v_g) \tag{2.29}$$

The presence of the v_g term, generally referred to as *ground-loop interference* or also *cross-talk for common return impedance,* may degrade the quality of the output signal appreciably, especially if v_i happens to be a low-level signal of magnitude comparable to v_g, as is often the case with transducer signals in industrial environments.

We can get rid of the v_g term by regarding v_i as a differential signal and v_g as a common-mode signal. Doing so requires changing the original amplifier to a difference-type amplifier and using an additional wire for direct access to the input

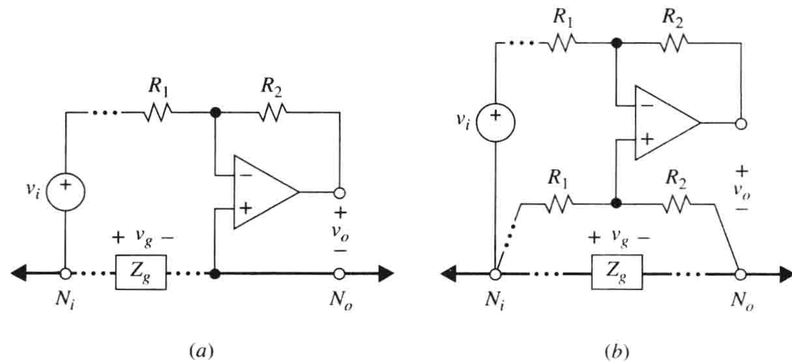


FIGURE 2.19
Using a difference amplifier to eliminate ground-loop interference.

signal common, in the manner shown in Fig. 2.19*b*. By inspection, we now have

$$v_O = -\frac{R_2}{R_1} v_i \tag{2.30}$$

The price we are paying in increased circuit complexity and wiring is certainly worth the benefits derived from the elimination of the v_g term.

2.5
INSTRUMENTATION AMPLIFIERS

An *instrumentation amplifier* (IA) is a difference amplifier meeting the following specifications: (*a*) extremely high (ideally infinite) common-mode and differential-mode input impedances; (*b*) very low (ideally zero) output impedance; (*c*) accurate and stable gain, typically in the range of 1 V/V to 10^3 V/V; and (*d*) extremely high common-mode rejection ratio. The IA is used to accurately amplify a low-level signal in the presence of a large common-mode component, such as a transducer output in process control and biomedicine. For this reason, IAs find widespread application in test and measurement instrumentation—hence the name.

With proper trimming, the difference amplifier of Fig. 2.13 can be made to meet the last three specifications satisfactorily. However, by Eq. (2.23), it fails to meet the first specification because both its differential-mode and its common-mode input resistances are finite; consequently, it will generally load down the circuit supplying the voltages v_1 and v_2, not to mention the ensuing degradation in the CMRR. These drawbacks are eliminated by preceding it with two high-input-impedance buffers. The result is a classic circuit known as the *triple-op-amp IA*.

Triple-Op-Amp IAs

In Fig. 2.20 OA_1 and OA_2 form what is often referred to as the *input* or *first* stage, and OA_3 forms the *output* or *second* stage. By the input voltage constraint, the voltage

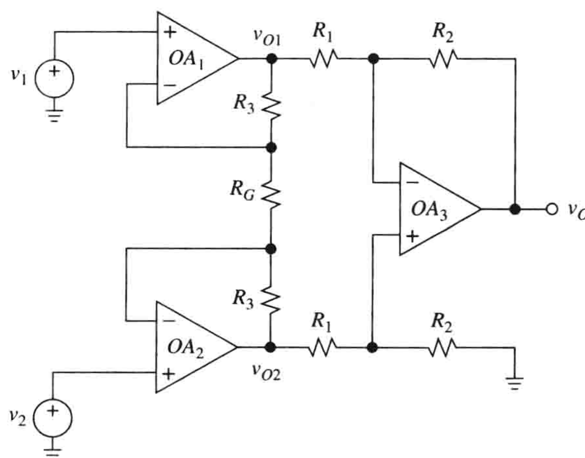


FIGURE 2.20
Triple-op-amp instrumentation amplifier.

across R_G is $v_1 - v_2$. By the input current constraint, the resistances denoted R_3 carry the same current as R_G. Applying Ohm's law yields $v_{O1} - v_{O2} = (R_3 + R_G + R_3)(v_1 - v_2)/R_G$, or

$$v_{O1} - v_{O2} = \left(1 + \frac{2R_3}{R_G}\right)(v_1 - v_2)$$

For obvious reasons the input stage is also referred to as a *difference-input, difference-output amplifier.* Next, we observe that OA_3 is a difference amplifier, and thus

$$v_O = \frac{R_2}{R_1}(v_{O2} - v_{O1})$$

Combining the last two equations gives

$$v_O = A(v_2 - v_1) \tag{2.31a}$$

$$A = A_I \times A_{II} = \left(1 + 2\frac{R_3}{R_G}\right) \times \left(\frac{R_2}{R_1}\right) \tag{2.31b}$$

indicating that the overall gain A is the product of the first- and second-stage gains A_I and A_{II}.

The gain depends on external resistance ratios, so it can be made quite accurate and stable by using resistors of suitable quality. Since OA_1 and OA_2 are operated in the noninverting configuration, their closed-loop input resistances are extremely high. Likewise, the closed-loop output resistance of OA_3 is quite low. Finally, the CMRR can be maximized by proper trimming of one of the second-stage resistances. We conclude that the circuit meets all the IA requisites listed earlier.

Equation (2.31b) points the way to go if variable gain is desired. To avoid perturbing bridge balance, we leave the second stage undisturbed and we vary gain by varying the single resistance R_G. If linear gain control is desired, we can use an arrangement of the type of Fig. 2.18.

EXAMPLE 2.10. (a) Design an IA whose gain can be varied over the range 1 V/V $\leq A \leq 10^3$ V/V by means of a 100-kΩ pot. (b) Make provisions for a trimmer to optimize its CMRR. (c) Outline a procedure for calibrating the trimmer.

Solution.

(a) Connect the 100-kΩ pot as a variable resistor, and use a series resistance R_4 to prevent R_G from going to zero. Since $A_I > 1$ V/V, we require $A_{II} < 1$ V/V in order to allow A to go all the way down to 1 V/V. Arbitrarily impose $A_{II} = R_2/R_1 = 0.5$ V/V, and use $R_1 = 100$ kΩ and $R_2 = 49.9$ kΩ, both 1%. By Eq. (2.31b), A_I must be variable from 2 V/V to 2000 V/V. At these extremes we have $2 = 1 + 2R_3/(R_4 + 100\text{ k}\Omega)$ and $2000 = 1 + 2R_3/(R_4 + 0)$. Solving, we obtain $R_4 = 50\ \Omega$ and $R_3 = 50$ kΩ. Use $R_4 = 49.9\ \Omega$ and $R_3 = 49.9$ kΩ, both 1%.

(b) Following Example 2.6, $4pR_2 = 4 \times 0.01 \times 49.9\text{ k}\Omega = 2$ kΩ. To be on the safe side, use a 47.5-kΩ, 1% resistor in series with a 5-kΩ pot. A suitable op amp is the OP-27 precision op amp (Analog Devices). The circuit is shown in Fig. 2.21.

(c) To calibrate the circuit, tie the inputs together and set the 100-kΩ pot for the maximum gain (wiper all the way up). Then, while switching the common inputs back and forth between −5 V and +5 V, adjust the 5-kΩ pot for the minimum change at the output.

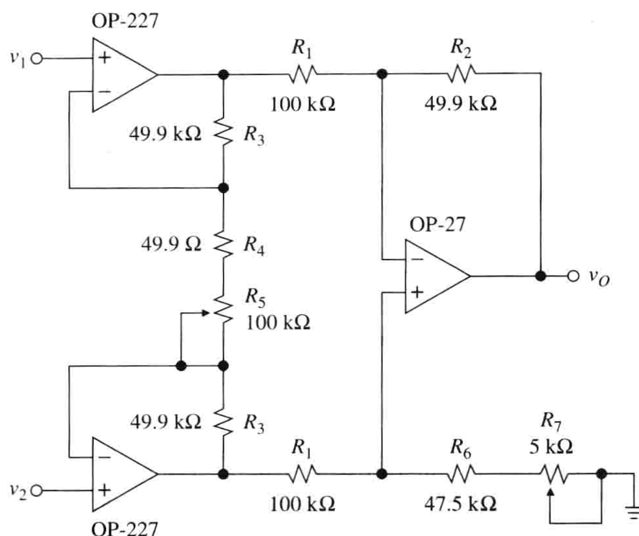


FIGURE 2.21
IA of Example 2.10.

The triple-op-amp IA configuration is available in IC form from various manufacturers. Familiar examples are the AD522 and INA101. These devices contain all components except for R_G, which is supplied externally by the user to set the gain, usually from 1 V/V to 10^3 V/V. Figure 2.22 shows a frequently used circuit symbol for the IA, along with its interconnection for remote sensing. In this arrangement, the *sense* and *reference* voltages are sensed right at the load terminals, so the effect of any signal losses in the long wires is eliminated by including these losses within the feedback loop. The accessibility to these terminals affords additional flexibility, such as the inclusion of an output power booster to drive high-current loads, or the offsetting of the output with respect to ground potential.

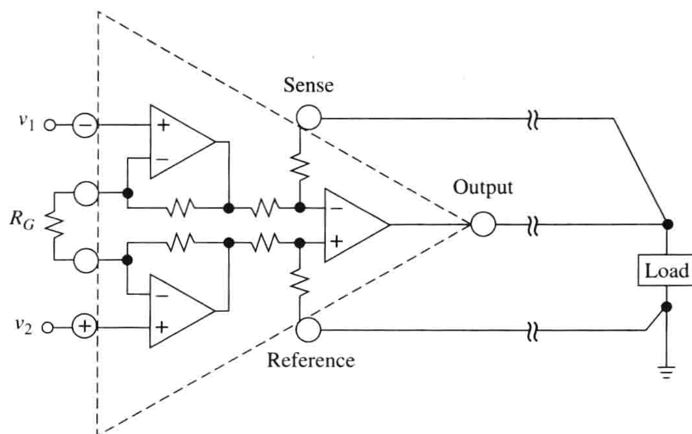


FIGURE 2.22
Standard IA symbol and connection for remote sensing.

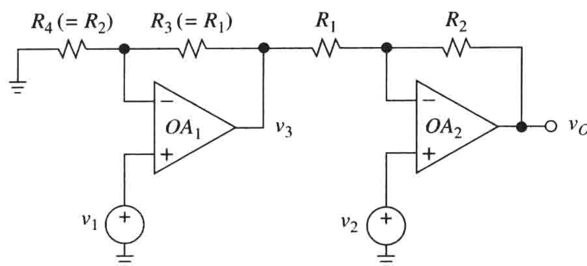


FIGURE 2.23
Dual-op-amp instrumentation amplifier.

Dual-Op-Amp IAs

When high-quality, costlier op amps are used to achieve superior performance, it is of interest to minimize the number of devices in the circuit. Shown in Fig. 2.23 is an IA that uses only two op amps. OA_1 is a noninverting amplifier, so $v_3 = (1 + R_3/R_4)v_1$. By the superposition principle, $v_O = -(R_2/R_1)v_3 + (1 + R_2/R_1)v_2$. Eliminating v_3, we can put v_O in the form

$$v_O = \left(1 + \frac{R_2}{R_1}\right) \times \left(v_2 - \frac{1 + R_3/R_4}{1 + R_1/R_2} v_1\right) \tag{2.32}$$

For true difference operation we require $1 + R_3/R_4 = 1 + R_1/R_2$, or

$$\frac{R_3}{R_4} = \frac{R_1}{R_2} \tag{2.33}$$

When this condition is met, we have

$$v_O = \left(1 + \frac{R_2}{R_1}\right)(v_2 - v_1) \tag{2.34}$$

Moreover, the circuit enjoys high input resistances and low output resistance. To maximize the CMRR, one of the resistors, say, R_4, should be trimmed. The adjustment of the trimmer proceeds as in the triple-op-amp case.

Adding a variable resistance between the inverting inputs of the two op amps as in Fig. 2.24 makes the gain adjustable. It can be shown (see Problem 2.45) that

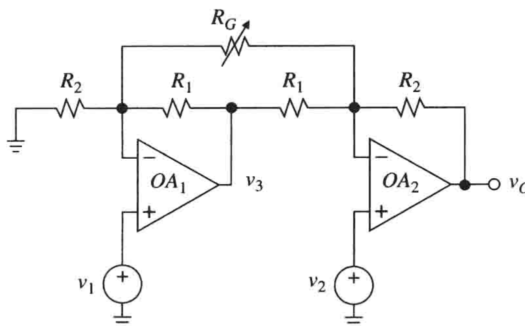


FIGURE 2.24
Dual-op-amp IA with variable gain.

$v_O = A(v_2 - v_1)$, where

$$A = 1 + \frac{R_2}{R_1} + \frac{2R_2}{R_G} \tag{2.35}$$

Compared with the triple-op-amp configuration, the dual-op-amp version offers the obvious advantage of requiring fewer resistors as well as one fewer op amp. The configuration is suited for realization with a dual-op-amp package, such as the OP227. The tighter matching usually available with dual op amps offers a significant boost in performance. A drawback of the dual-op-amp configuration is that it treats the inputs asymmetrically because v_1 has to propagate through OA_1 before catching up with v_2. Because of this additional delay, the common-mode components of the two signals will no longer cancel each other out as frequency is increased, leading to a premature degradation of the CMRR with frequency. Conversely, the triple-op-amp configuration enjoys a higher degree of symmetry and usually maintains high CMRR performance over a broader frequency range. The factors limiting the CMRR here are mismatches in the delays through the first-stage op amps, as well as bridge imbalance and common-mode limitations of the second-stage op amp.

Monolithic IAs

The need for instrumentation amplification arises so often that it justifies the manufacture of special ICs to perform just this function.[2] Compared with realizations built using general-purpose op amps, this approach allows better optimization of the parameters that are critical to this application, particularly the CMRR, gain linearity, and noise.

The task of first-stage difference amplification as well as common-mode rejection is delegated to highly matched transistor pairs. A transistor pair is faster than a pair of full-fledged op amps and can be made to be less sensitive to common-mode signals, thus relaxing the need for very tightly matched resistances. Examples of dedicated IC IAs are the AD521/524/624/625 and the AMP01 and AMP05.

Figure 2.25 shows a simplified circuit diagram of the AMP01, and Fig. 2.26 shows the basic interconnection to make it work with gains ranging from 0.1 V/V to 10^4 V/V. As shown, the gain is set by the ratio of two user-supplied resistors R_S and R_G as

$$A = 20\frac{R_S}{R_G} \tag{2.36}$$

With this arrangement one can achieve highly stable gains by using a pair of temperature-tracking resistors.

Referring to Fig. 2.25 and the connection of Fig. 2.26, we can describe circuit operation as follows. Applying a differential signal between the inputs unbalances the currents through Q_1 and Q_2. A_1 reacts to this by unbalancing Q_1 and Q_2 in the opposite direction in order to restore the balanced condition $v_N = v_P$ at its own inputs. A_1 achieves this by applying a suitable drive to the bottom transistor pair via A_3. The amount of drive needed depends on the ratio R_S/R_G as well as on the magnitude of the input difference. This drive forms the output of the IA. Searching the web for the AMP01 data sheets, you will find that $\text{CMRR}_{\text{dB}} = 140$ dB at $A = 10^4$ V/V.

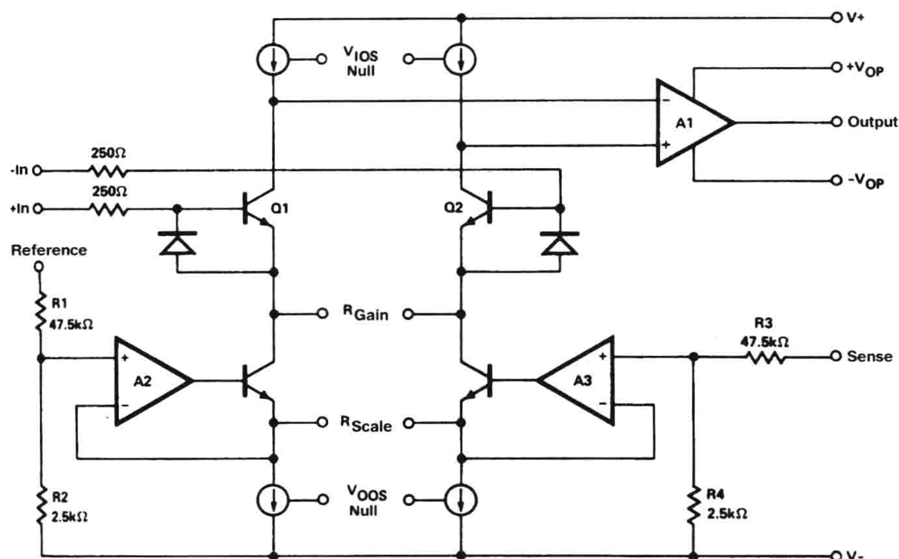


FIGURE 2.25
Simplified circuit diagram of the AMP01 low-noise precision IA. (Courtesy of Analog Devices.)

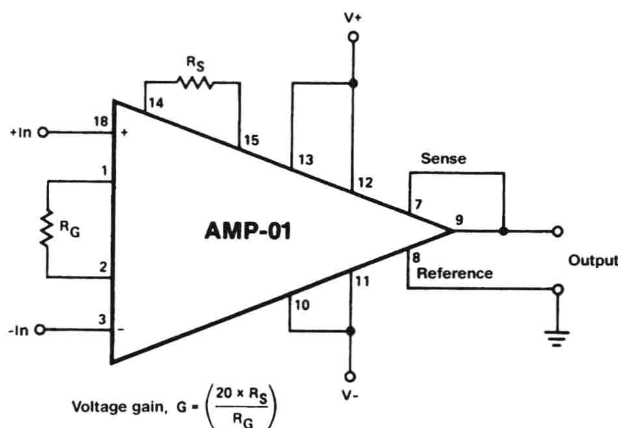


FIGURE 2.26
Basic AMP01 connection for gains from 0.1 V/V to 10 V/mV. (Courtesy of Analog Devices.)

Flying-Capacitor Techniques

A popular alternative for achieving high CMRRs is the *flying-capacitor technique,* so called because it flips a capacitor back and forth between source and amplifier. As exemplified[3] in Fig. 2.27, flipping the switches to the left charges C_1 to the voltage difference $v_2 - v_1$, and flipping the switches to the right transfers charge from C_1 to C_2. Continuous switch clocking causes C_2 to charge up until the equilibrium

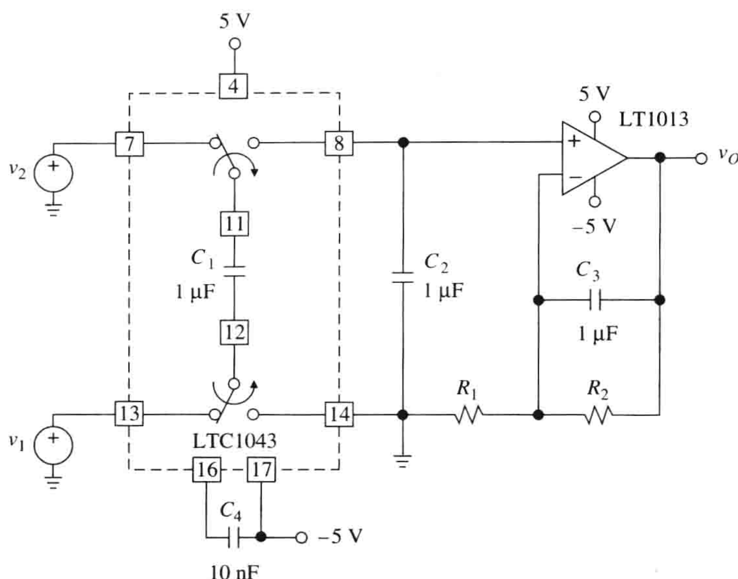


FIGURE 2.27
Flying-capacitor IA. (Courtesy of Linear Technology.)

condition is reached in which the voltage across C_2 becomes equal to that across C_1. This voltage is magnified by the noninverting amplifier to give

$$v_O = \left(1 + \frac{R_2}{R_1}\right)(v_2 - v_1) \tag{2.37}$$

To achieve high performance, the circuit shown uses the LTC1043 precision instrumentation switched-capacitor building block and the LT1013 precision op amp. The former includes an on-chip clock generator to operate the switches at a frequency set by C_4. With $C_4 = 10$ nF, this frequency is 500 Hz. The function of C_3 is to provide low-pass filtering to ensure a clean output. Thanks to the flying-capacitor technique, the circuit completely ignores common-mode input signals to achieve a high CMRR, typically[3] in excess of 120 dB at 60 Hz.

2.6 INSTRUMENTATION APPLICATIONS

In this section we examine some issues arising in the application of instrumentational amplifiers.[2,4] Additional applications will be discussed in the next section.

Active Guard Drive

In applications such as the monitoring of hazardous industrial conditions, source and amplifier may be located far apart from each other. To help reduce the effect of noise pickup as well as ground-loop interference, the input signal is transmitted in double-ended form over a pair of shielded wires and then processed with a difference

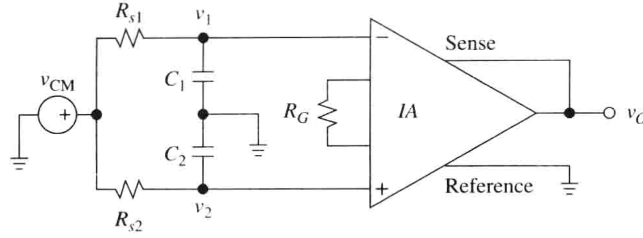


FIGURE 2.28
Model of nonzero source resistance and distributed cable capacitance.

amplifier, such as an IA. The advantage of double-ended over single-ended transmission is that since the two wires tend to pick up identical noise, this noise will appear as a common-mode component and will thus be rejected by the IA. For this reason, double-ended transmission is also referred to as *balanced transmission.* The purpose of shielding is to help reduce differential-mode noise pickup.

Unfortunately, because of the distributed capacitance of the cable, another problem arises, namely, CMRR degradation with frequency. To investigate this aspect, refer to Fig. 2.28, where the source resistances and cable capacitances have been shown explicitly. Since the differential-mode component has been assumed to be zero, we expect the output of the IA to be likewise zero. In practice, since the time constants $R_{s1}C_1$ and $R_{s2}C_2$ are likely to be different, any variation in v_{CM} will produce uneven signal variations downstream of the RC networks, or $v_1 \neq v_2$, thus resulting in a differential error signal that the IA will then amplify and reproduce at the output. The effect of RC imbalance is therefore a nonzero output signal in spite of the absence of any differential-mode component at the source. This represents a degradation in the CMRR.

The CMRR due to RC imbalance is

$$\text{CMRR}_{\text{dB}} \cong 20 \log_{10} \frac{1}{2\pi f R_{\text{dm}} C_{\text{cm}}} \tag{2.38}$$

where $R_{\text{dm}} = |R_{s1} - R_{s2}|$ is the source resistance imbalance, $C_{\text{cm}} = (C_1 + C_2)/2$ is the common-mode capacitance between each wire and the grounded shield, and f is the frequency of the common-mode input component. For instance, at 60 Hz, a source resistance imbalance of 1 kΩ in conjunction with a 100-foot cable having a distributed capacitance of 1 nF would degrade the CMRR to $20 \log_{10}[1/(2\pi 60 \times 10^3 \times 10^{-9})] = 68.5$ dB, even with an IA having infinite CMRR.

The effect of C_{cm} can, to a first approximation, be neutralized by driving the shield with the common-mode voltage itself so as to reduce the common-mode swing across C_{cm} to zero. Figure 2.29 shows a popular way of achieving this goal. By op amp action, the voltages at the top and bottom nodes of R_G are v_1 and v_2. Denoting the voltage across R_3 as v_3, we can write $v_{\text{CM}} = (v_1 + v_2)/2 = (v_1 + v_3 + v_2 - v_3)/2 = (v_{O1} + v_{O2})/2$, indicating that v_{CM} can be extracted by computing the mean of v_{O1} and v_{O2}. This mean is found via the two 20-kΩ resistors and is then buffered to the shield by OA_4.

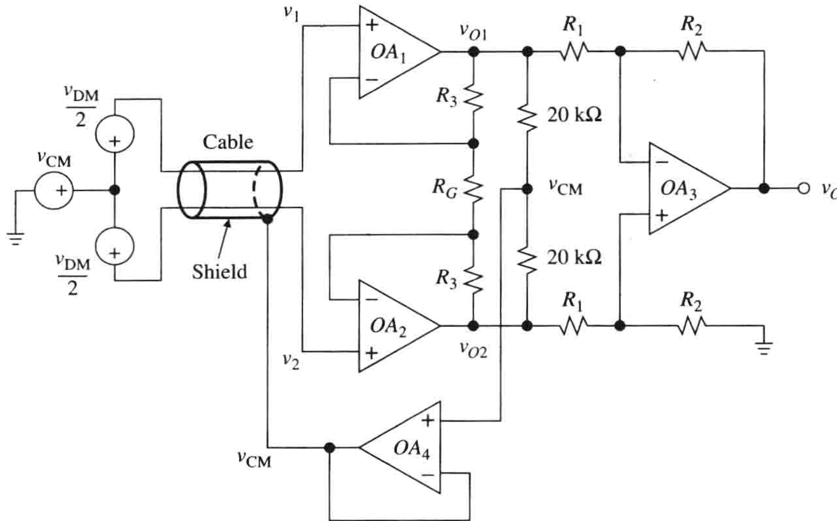


FIGURE 2.29
IA with active guard drive.

Digitally Programmable Gain

In automatic instrumentation, such as data acquisition systems, it is often desirable to program the gain of the IA electronically, usually by means of JFET or MOSFET switches. The method depicted in Fig. 2.30 programs the first-stage gain A_I by using a string of symmetrically valued resistors, and a string of simultaneously activated switch pairs to select the tap pair corresponding to a given gain. At any given time, only one switch pair is closed and all others are open. By Eq. (2.31*b*), A_I can be put in the form

$$A_I = 1 + \frac{R_{\text{outside}}}{R_{\text{inside}}} \tag{2.39}$$

where R_{inside} is the sum of the resistances located between the two selected switches and R_{outside} is the sum of all remaining resistances. For the case shown, the selected switch pair is SW_1, so $R_{\text{outside}} = 2R_1$ and $R_{\text{inside}} = 2(R_2 + R_3 + \cdots + R_n) + R_{n+1}$. Selecting SW_2 gives $R_{\text{outside}} = 2(R_1 + R_2)$ and $R_{\text{inside}} = 2(R_3 + \cdots + R_n) + R_{n+1}$. It is apparent that changing to a different switch pair increases (or decreases) R_{outside} at the expense of an equal decrease (or increase) in R_{inside}, thus yielding a different resistance ratio and, hence, a different gain.

The advantage of this topology is that the current flowing through any closed switch is the negligible input current of the corresponding op amp. This is particularly important when the switches are implemented with FETs because FETs have a nonzero on-resistance, and the ensuing voltage drop could degrade the accuracy of the IA. With zero current this drop is also zero, in spite of the nonideality of the switch.

The two groups of switches in Fig. 2.30 can easily be implemented with CMOS analog multiplexers/demultiplexers, such as the CD4051 or CD4052.

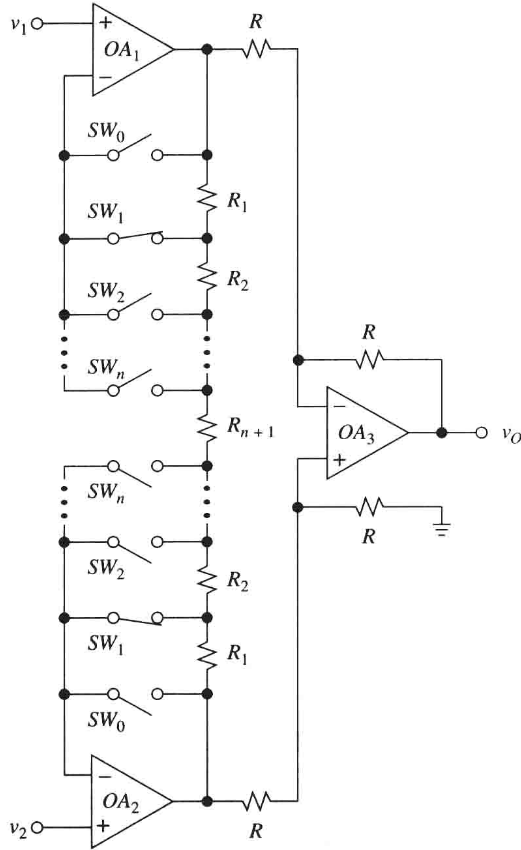


FIGURE 2.30
Digitally programmable IA.

Digitally programmable IAs, containing all the necessary resistors, analog switches, and TTL-compatible decoder and switch-driver circuitry, are also available in IC form. Consult the manufacturer catalogs for more information.

Output-Offsetting

There are applications that call for a prescribed amount of offset at the output of an IA, as when an IA is fed to a voltage-to-frequency converter, which requires that its input range be of only one polarity. Since the IA output is usually bipolar, it must be suitably offset to ensure a unipolar range. In the circuit of Fig. 2.31 the reference node is driven by voltage V_{REF}. This voltage, in turn, is obtained from the wiper of a pot and is buffered by OA_4, whose low output resistance prevents disturbance of the bridge balance. Applying the superposition principle, we obtain $v_O = A(v_2 - v_1) + (1 + R_2/R_1) \times [R_1/(R_1 + R_2)]V_{REF}$, or

$$v_O = A(v_2 - v_1) + V_{REF} \tag{2.40}$$

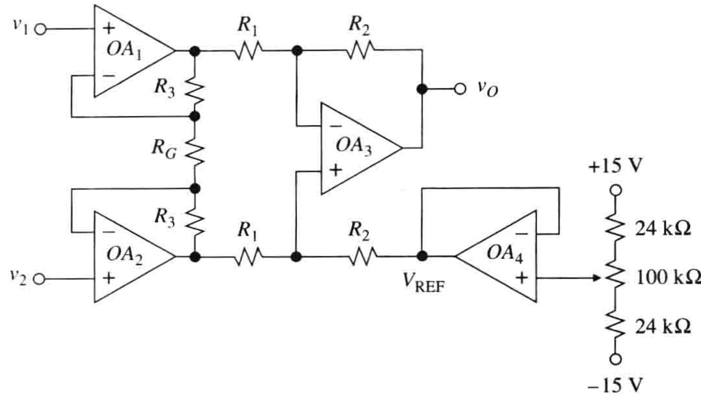


FIGURE 2.31
IA with output offset control.

where A is given by Eq. (2.31b). With the component values shown, V_{REF} is variable from -10 V to $+10$ V.

Current-Output IAs

By turning the second stage into a Howland circuit, in the manner depicted in Fig. 2.32, we can configure the triple-op-amp IA for current-output operation. This type of operation is desirable when transmitting signals over long wires since the stray wire resistance does not degrade current signals. Combining the results of Problem 2.13 with Eq. (2.31b), we readily obtain

$$i_O = \frac{1 + 2R_3/R_G}{R_1}(v_2 - v_1) \tag{2.41}$$

The gain can be adjusted via R_G, as usual. For efficient operation, the Howland stage can be improved with the modification of Fig. 2.11. For high CMRR, the top left resistance should be trimmed.

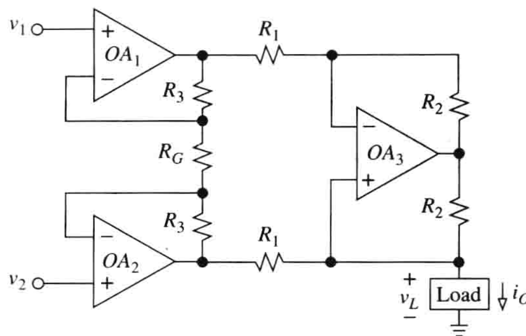


FIGURE 2.32
Current-output IA.

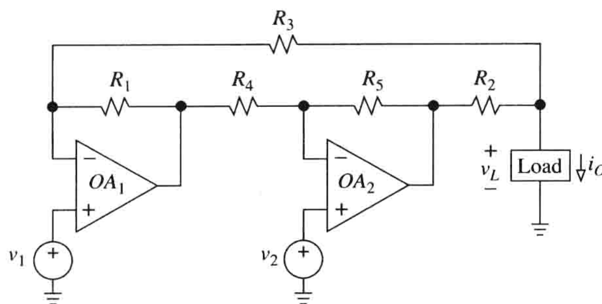


FIGURE 2.33
Dual-op-amp IA with current output.

The dual-op-amp IA is configured for current-output operation by the bootstrapping technique of Fig. 2.33. It is left as an exercise (see Problem 2.52) to prove that the transfer characteristic of the circuit is of the type

$$i_O = \frac{1}{R}(v_2 - v_1) - \frac{1}{R_o}v_L \tag{2.42a}$$

$$R_o = \frac{R_2/R_1}{R_5/R_4 - (R_2 + R_3)/R_1}R_3 \tag{2.42b}$$

so that imposing $R_2 + R_3 = R_1 R_5/R_4$ yields $R_o = \infty$. If adjustable gain is desired, it is readily obtained by connecting a variable resistance R_G between the inverting input pins of the two op amps, in the manner of Fig. 2.24.

Besides offering difference-input operation with high input resistances, the circuit enjoys the efficiency advantages of the improved Howland circuit because R_2 can be kept as small as needed while all remaining resistances can be made relatively large to conserve power. When this constraint is imposed, the voltage compliance is approximately $|v_L| \leq V_{\text{sat}} - R_2|i_O| = V_{\text{sat}} - 2|v_2 - v_1|$.

Current-Input IA

In current-loop instrumentation the need arises for sensing a floating current and converting it to a voltage. To avoid perturbing the characteristics of the loop, it is desirable that the circuit downstream appear as a virtual short. An IA can once again be suitably modified to meet this requirement. In Fig. 2.34 we observe that OA_1 and OA_2 force the voltages at their input pins to track v_{CM}, thus ensuring 0 V across the input source. By KVL and Ohm's law, $v_{O2} = v_{CM} - R_3 i_I$ and $v_{O1} = v_{CM} + R_3 i_I$. But $v_O = (R_2/R_1) \times (v_{O2} - v_{O1})$. Combining, we get

$$v_O = -\frac{2R_2}{R_1}R_3 i_I \tag{2.43}$$

If variable gain is desired, this can be obtained by modifying the difference stage as in Fig. 2.17 or 2.18. If, on the other hand, the difference stage is modified as in Fig. 2.32, the circuit becomes a floating-input current amplifier.

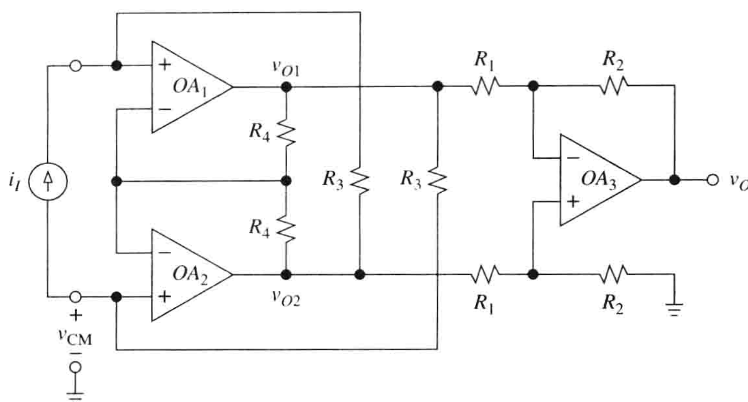


FIGURE 2.34
Current-input IA.

2.7 TRANSDUCER BRIDGE AMPLIFIERS

Resistive transducers are devices whose resistance varies as a consequence of some environmental condition, such as temperature (thermistors; resistance temperature detectors, or RTDs), light (photoresistors), strain (strain gauges), and pressure (piezoresistive transducers). By making these devices part of a circuit, it is possible to produce an electric signal that, after suitable conditioning, can be used to monitor as well as control the physical process affecting the transducer.[5] In general it is desirable that the relationship between the final signal and the original physical variable be linear, so that the former can directly be calibrated in the physical units of the latter. Transducers play such an important role in measurement and control instrumentation that it is worth studying transducer circuits in some detail.

Transducer Resistance Deviation

Transducer resistances are expressed in the form $R + \Delta R$, where R is the resistance at some *reference* condition, such as 0°C in the case of temperature transducers, or the absence of strain in the case of strain gauges, and ΔR represents the *deviation* from the reference value as a consequence of a change in the physical condition affecting the transducer. Transducer resistances are also expressed in the alternative form $R(1 + \delta)$, where $\delta = \Delta R/R$ represents the *fractional deviation.* Multiplying δ by 100 yields the *percentage deviation.*

EXAMPLE 2.11. Platinum resistance temperature detectors (Pt RTDs) have a temperature coefficient[6] $\alpha = 0.00392/°\text{C}$. A popular Pt RTD reference value at $T = 0$ °C is 100 Ω. (*a*) Write an expression for the resistance as a function of T. (*b*) Compute $R(T)$ for $T = 25$ °C, 100 °C, −15 °C. (*c*) Calculate ΔR and δ for a temperature change $\Delta T = 10$ °C.

Solution.

(*a*) $R(T) = R(0\,°\text{C})(1 + \alpha T) = 100(1 + 0.00392T)\ \Omega$.
(*b*) $R(25\,°\text{C}) = 100(1 + 0.00392 \times 25) = 109.8\ \Omega$. Likewise, $R(100\,°\text{C}) = 139.2\ \Omega$ and $R(-15\,°\text{C}) = 94.12\ \Omega$.
(*c*) $R + \Delta R = 100 + 100\alpha T = 100 + 100 \times 0.00392 \times 10 = 100\ \Omega + 3.92\ \Omega$; $\delta = \alpha\Delta T = 0.00392 \times 10 = 0.0392$. This corresponds to a change of $0.0392 \times 100 = 3.92\%$.

The Transducer Bridge

To measure resistance deviation, we must find a method to convert ΔR to a voltage variation ΔV. The simplest technique is to make the transducer part of a voltage divider, as shown in Fig. 2.35. The transducer voltage is $v_1 = V_{\text{REF}} R(1 + \delta)/[R_1 + R(1 + \delta)]$, which can be put in the insightful form

$$v_1 = \frac{R}{R_1 + R} V_{\text{REF}} + \frac{\delta V_{\text{REF}}}{2 + R_1/R + R/R_1 + (1 + R/R_1)\delta} \tag{2.44}$$

where $\delta = \Delta R/R$. We observe that v_1 consists of a fixed term plus a term controlled by $\delta = \Delta R/R$. It is precisely the latter that interests us, so we must find a means for amplifying it while ignoring the former. This is achieved by using a second voltage divider to synthesize the term

$$v_2 = \frac{R}{R_1 + R} V_{\text{REF}} \tag{2.45}$$

and then using an IA to take the difference $v_1 - v_2$. Denoting the IA gain as A, we get $v_O = A(v_1 - v_2)$, or

$$v_O = A V_{\text{REF}} \frac{\delta}{1 + R_1/R + (1 + R/R_1)(1 + \delta)} \tag{2.46}$$

The four-resistor structure is the familiar resistive bridge, and the two voltage dividers are referred to as the *bridge legs.*

It is apparent that v_O is a nonlinear function of δ. In microprocessor-based systems, a nonlinear function can easily be linearized in the software. Quite often, however, we have $\delta \ll 1$, so

$$v_O \cong \frac{A V_{\text{REF}}}{2 + R_1/R + R/R_1} \delta \tag{2.47}$$

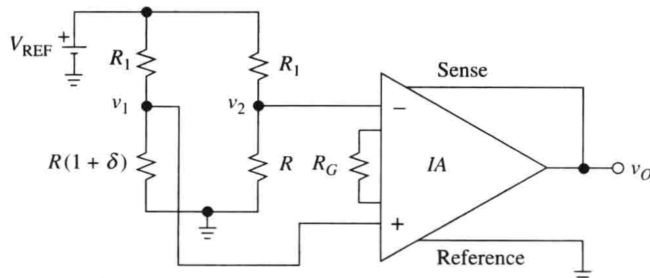


FIGURE 2.35
Transducer bridge and IA.

indicating a linear dependence of v_O on δ. Many bridges are designed with $R_1 = R$, in which case Eqs. (2.46) and (2.47) become

$$v_O = \frac{AV_{REF}}{4}\frac{\delta}{1+\delta/2} \tag{2.48}$$

$$v_O \cong \frac{AV_{REF}}{4}\delta \tag{2.49}$$

EXAMPLE 2.12. Let the tranducer of Fig. 2.35 be the Pt RTD of Example 2.11, and let $V_{REF} = 15$ V. (*a*) Specify values for R_1 and A suitable for achieving an output sensitivity of 0.1 V/°C near 0 °C. To avoid self-heating in the RTD, limit its power dissipation to less than 0.2 mW. (*b*) Compute $v_O(100\ °C)$ and estimate the equivalent error, in degrees Celsius, in making the approximation of Eq. (2.47).

Solution.

(*a*) Denoting the transducer current as i, we have $P_{RTD} = Ri^2$. Thus, $i^2 \leq P_{RTD(max)}/R = 0.2 \times 10^{-3}/100$, or $i = 1.41$ mA. To be on the safe side, impose $i \cong 1$ mA, or $R_1 = 15$ kΩ. For $\Delta T = 1$ °C we have $\delta = \alpha \times 1 = 0.00392$, and we want $\Delta v_O = 0.1$ V. By Eq. (2.47) we need $0.1 = A \times 15 \times 0.00392/(2+15/0.1+0.1/15)$, or $A = 258.5$ V/V.

(*b*) For $\Delta T = 100$ °C we have $\delta = \alpha\Delta T = 0.392$. Inserting into Eq. (2.46), we get $v_O(100\ °C) = 9.974$ V. Equation (2.47) predicts that $v_O(100\ °C) = 10.0$ V, which exceeds the actual value by $10 - 9.974 = 0.026$ V. Since 0.1 V corresponds to 1 °C, 0.026 V corresponds to $0.026/0.1 = 0.26$ °C. Therefore, in using the approximated expression, we cause, at 100 °C, an error of about one-quarter of a degree Celsius.

Bridge Calibration

With $\Delta R = 0$, a transducer bridge should be balanced and yield a zero voltage difference between its taps. In practice, because of resistance tolerances, including the tolerance of the transducer's reference value, the bridge is likely to be unbalanced and a trimmer should be included to balance it. Moreover, the tolerances in the values of the resistances and of V_{REF} will affect the bridge sensitivity $(v_1 - v_2)/\delta$, thus creating the need for adjustment of this parameter as well.

Figure 2.36 shows a circuit that allows for both adjustments. Varying R_2's wiper from its midway position will assign more resistance to one leg and less to the other,

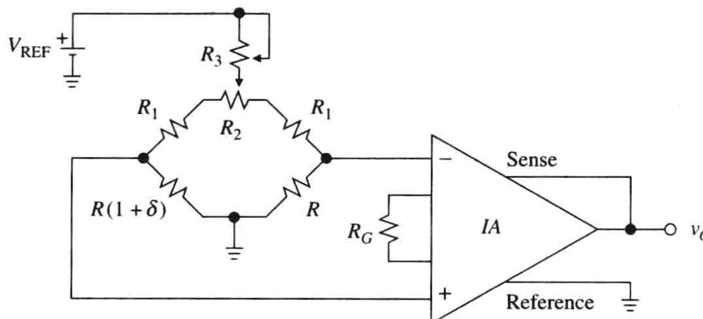


FIGURE 2.36
Bridge calibration.

thus allowing the compensation of their inherent mismatches. Varying R_3 changes the bridge current and hence the magnitude of the voltage variation produced by the transducer, thus allowing the adjustment of the sensitivity.

EXAMPLE 2.13. Let all resistors in Example 2.12 have a 1% tolerance, and let V_{REF} have a 5% tolerance. (*a*) Design a circuit to calibrate the bridge. (*b*) Outline the calibration procedure.

Solution.

(*a*) A 5% tolerance in V_{REF} means that its actual value can deviate from its nominal value by as much as $\pm 0.05 \times 15 = \pm 0.75$ V. To be on the safe side and to also include the effect of 1% resistance tolerance, assume a maximum deviation of ± 1 V, and thus design for 14 V $\pm$ 1 V at R_2's wiper. To ensure a current of 1 mA at each leg, we need $R_3 = 2/(1+1) = 1$ kΩ and $R + R_1 + R_2/2 = 14/1 = 14$ kΩ. Since R_2 must compensate for up to a 1% variation on each leg, we need $R_2 = 2 \times 0.01 \times 14$ k$\Omega = 280$ Ω. To be on the safe side, pick $R_2 = 500$ Ω. Then $R_1 = 14$ k$\Omega - 100$ $\Omega - 500/2$ $\Omega = 13.65$ kΩ (use 13.7 kΩ, 1%). The IA gain A must be recomputed via Eq. (2.47), but with $V_{REF} = 14$ V and with 13.7 k$\Omega + 500/2$ $\Omega = 13.95$ kΩ in place of R_1. This yields $A = 257.8$ V/V. Summarizing, we need $R_1 = 13.7$ kΩ, 1%; $R_2 = 500$ Ω; $R_3 = 1$ kΩ; and $A = 257.8$ V/V.

(*b*) To calibrate, first set $T = 0$ °C and adjust R_2 for $v_O = 0$ V. Then set $T = 100$ °C and adjust R_3 for $v_O = 10.0$ V.

Strain-Gauge Bridges

The resistance of a wire having resistivity ρ, cross-sectional area S, and length ℓ is $R = \rho\ell/S$. Straining the wire changes its length to $\ell + \Delta\ell$, its area to $S - \Delta S$, and its resistance to $R + \Delta R = \rho(\ell + \Delta\ell)/(S - \Delta S)$. Since its volume must remain constant, we have $(\ell + \Delta\ell) \times (S - \Delta S) = S\ell$. Eliminating $S - \Delta S$, we get $\Delta R = R(\Delta\ell/\ell)(2 + \Delta\ell/\ell)$. But $\Delta\ell/\ell \ll 2$, so

$$\Delta R = 2R\frac{\Delta\ell}{\ell} \tag{2.50}$$

where R is the *unstrained* resistance and $\Delta\ell/\ell$ is the *fractional elongation.* A strain gauge is fabricated by depositing resistive material on a flexible backing according to a pattern designed to maximize its fractional elongation for a given strain. Since strain gauges are sensitive also to temperature, special precautions must be taken to mask out temperature-induced variations. A common solution is to work with gauge pairs designed to compensate for each other's temperature variations.

The strain-gauge arrangement of Fig. 2.37 is referred to as a *load cell.* Denoting the bridge voltage as V_B and ignoring R_1 for a moment, the voltage divider formula yields $v_1 = V_B(R + \Delta R)/(R + \Delta R + R - \Delta R) = V_B(R + \Delta R)/2R$, $v_2 = V_B(R - \Delta R)/2R$, and $v_1 - v_2 = V_B\Delta R/R = V_B\delta$, so

$$v_O = AV_{REF}\delta \tag{2.51}$$

The sensitivity is now four times as large as that given in Eq. (2.49), thus relaxing the demands upon the IA. Furthermore, the dependence of v_O on δ is now perfectly linear—another advantage of working with gauge pairs. To achieve the $+\Delta R$ and

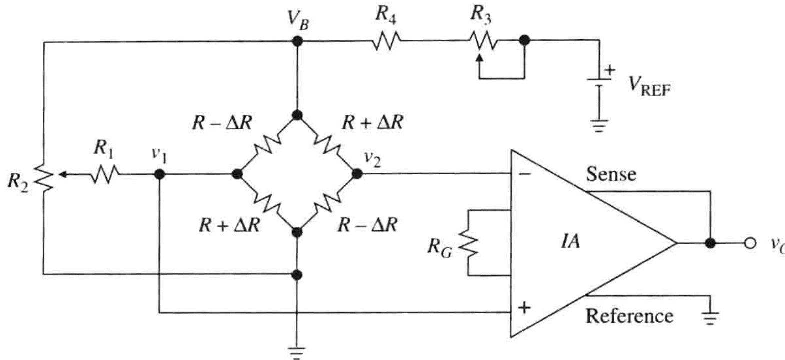


FIGURE 2.37
Strain-gauge bridge and IA.

$-\Delta R$ variations, two of the gauges will be bonded to one side of the structure under strain, and the other two to the opposite side. Even in installations in which only one side is accessible, it pays to work with four gauges because two can be used as dummy gauges to provide temperature compensation for the active ones. Piezoresistive pressure sensors also use this arrangement.

Figure 2.37 also illustrates an alternative technique for balancing the bridge. In the absence of strain, each tap voltage should be $V_B/2$. In practice there will be deviations due to the initial tolerances of the four gauges. By varying R_2's wiper, we can force an adjustable amount of current through R_1 that will increase or decrease the corresponding tap voltage until the bridge is nulled. Resistors R_3 and R_4 drop V_{REF} to V_B, and R_3 adjusts the sensitivity.

EXAMPLE 2.14. Let the strain gauges of Fig. 2.37 be 120-Ω, ±1% types, and let their maximum current be limited to 20 mA to avoid excessive self-heating. (*a*) Assuming that $V_{REF} = 15\text{ V} \pm 5\%$, specify suitable values for R_1 through R_4. (*b*) Outline the calibration procedure.

Solution.

(*a*) By Ohm's law, $V_B = 2 \times 120 \times 20 \times 10^{-3} = 4.8$ V. In the absence of strain, the tap voltages are nominally $V_B/2 = 2.4$ V. Their actual values may deviate from $V_B/2$ by as much as ±1% of 2.4 V, that is, by as much as ±0.024 V. Consider the case in which $v_1 = 2.424$ V and $v_2 = 2.376$ V. By moving R_2's wiper to ground, we must be able to lower v_1 to 2.376 V, that is, to change v_1 by 0.048 V. To achieve this, R_1 must sink a current $i = 0.048/(120 \parallel 120) = 0.8$ mA, so $R_1 \cong 2.4/0.8 = 3$ kΩ (to be on the safe side, use $R_1 = 2.37$ kΩ, 1%). To prevent excessive loading of R_2's wiper by R_1, use $R_2 = 1$ kΩ. Under nominal conditions we have $i_{R_3} = i_{R_4} = 2 \times 20 \times 10^{-3} + 4.8/10^3 \cong 45$ mA. Following Example 2.13, we wish R_3 to drop a maximum of 2 V. So $R_3 = 2/45 = 44$ Ω (use $R_3 = 50$ Ω). With R_3's wiper halfway we have $R_4 = (15 - 25 \times 45 \times 10^{-3} - 4.8)/(45 \times 10^{-3}) = 202$ Ω (use 200 Ω). Summarizing, $R_1 = 2.37$ kΩ, $R_2 = 1$ kΩ, $R_3 = 50$ Ω, and $R_4 = 200$ Ω.

(*b*) To calibrate, first adjust R_2 so that with no strain we get $v_O = 0$ V. Then apply a known strain, preferably near the full scale, and adjust R_3 for the desired value of v_O.

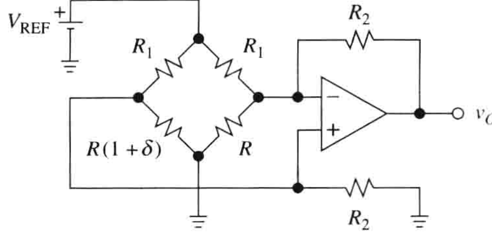


FIGURE 2.38
Single-op-amp bridge amplifier.

Single-Op-Amp Amplifier

For reasons of cost it is sometimes desirable to use a simpler amplifier than the full-fledged IA. Figure 2.38 shows a bridge amplifier implemented with a single op amp. After applying Thévenin's theorem to the two legs of the bridge, we end up with the familiar difference amplifier. One can then show (see Problem 2.57) that

$$v_O = \frac{R_2}{R} V_{\text{REF}} \frac{\delta}{R_1/R + (1 + R_1/R_2)(1 + \delta)} \tag{2.52}$$

For $\delta \ll 1$ this simplifies to

$$v_O \cong \frac{R_2}{R} V_{\text{REF}} \frac{\delta}{1 + R_1/R + R_1/R_2} \tag{2.53}$$

That is, v_O depends linearly on δ. To adjust the sensitivity and to null the effect of resistance mismatches, we can use a scheme of the type of Fig. 2.36.

Bridge Linearization

With the exception of the strain-gauge circuit of Fig. 2.37, all bridge circuits discussed so far suffer from the fact that the response is reasonably linear only as long as $\delta \ll 1$. It is therefore of interest to seek circuit solutions capable of a linear response regardless of the magnitude of δ.

The design of Fig. 2.39 linearizes the bridge by driving it with a constant current.[7] This is achieved by placing the entire bridge within the feedback loop

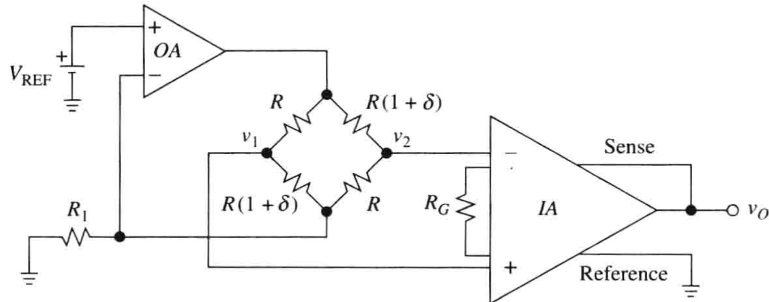


FIGURE 2.39
Bridge linearization by constant-current drive.

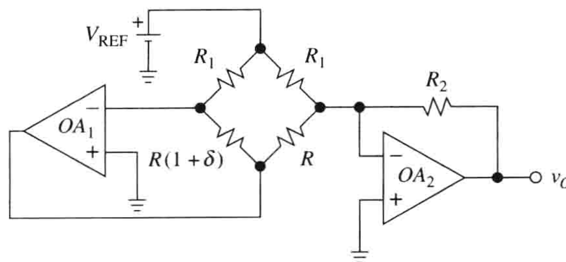


FIGURE 2.40
Single-transducer circuit with linear response.

of a floating-load V-I converter. The bridge current is $I_B = V_{REF}/R_1$. By using a transducer pair as shown, I_B will split equally between the two legs. Since OA keeps the bottom node of the bridge at V_{REF}, we have $v_1 = V_{REF} + R(1+\delta)I_B/2$, $v_2 = V_{REF} + RI_B/2$, and $v_1 - v_2 = R\delta I_B/2$, so

$$v_O = \frac{ARV_{REF}}{2R_1}\delta \tag{2.54}$$

The alternative design of Fig. 2.40 uses a single-transducer element and a pair of inverting-type op amps.[7] The response is again linearized by placing the bridge within the feedback loop of the V-I converter OA_1. It is left as an exercise (see Problem 2.57) to show that

$$v_O = \frac{R_2 V_{REF}}{R_1}\delta \tag{2.55}$$

For additional bridge circuit examples, see references 5, 7, and 8 and the end-of-chapter problems.

PROBLEMS

2.1 Current-to-voltage converters

2.1 (*a*) Using an op amp powered from ±10-V regulated supplies, design a circuit that accepts a source i_S having parallel resistance R_s, and yields an output voltage v_O such that as i_S is varied from 0 to +1 mA (flowing into your circuit), v_O varies from +5 V to −5 V. *Hint:* You need to offset the output. (*b*) Find the minimum gain a for a closed-loop gain deviation from the ideal by less than 0.01% for $R_s = \infty$. (*c*) Find the minimum value of R_s for a closed-loop gain departure from the ideal by less than 0.025% with the minimum value of a found in (*b*).

2.2 (*a*) Using two op amps, design a circuit that accepts two input sources i_{S1} and i_{S2}, both flowing into your circuit, and having parallel resistance R_{s1} and R_{s2}, and gives, for $a \to \infty$, $v_O = A_1 i_{S1} - A_2 i_{S2}$, where $A_1 = A_2 =$ (10 V/mA). (*b*) How are A_1 and A_2 affected if $R_{s1} = R_{s2} = 30$ kΩ and the op amps have $a = 10^3$ V/V?

2.3 Design a circuit to convert a 4-mA-to-20-mA input current to a 0-V-to-8-V output voltage. The reference direction of the input source is from ground into your circuit, and the circuit is powered from ±10-V regulated supplies.

2.4 Estimate the closed-loop parameters if the circuit of Example 2.2 is implemented with a 741 op amp.

2.5 (*a*) Using an op amp powered from ±5-V regulated supplies, design a photodetector amplifier such that as the photodiode current varies from 0 to 10 μA, v_O varies from -4 V to $+4$ V, under the constraint that no resistance greater than 100 kΩ be used. (*b*) Find the minimum gain a for a less than 0.1% closed-loop gain deviation from the ideal.

2.6 (*a*) Let the op amp of Fig. P1.65 have $a = 10^4$ V/V, and let $R_1 = R_3 = R_5 = 20$ kΩ and $R_2 = R_4 = 10$ kΩ. Find v_O/i_I as well as the resistance R_i seen by the input source.

2.2 Voltage-to-current converters

2.7 (*a*) Show that the floating-load V-I converter of Fig. P2.7 yields $i_O = v_I/(R_1/k)$, $k = 1 + R_2/R_3$. (*b*) Specify standard 5% resistances for a sensitivity of 2 mA/V and $R_i = 1$ MΩ, where R_i is the resistance seen by the input source. (*c*) If $\pm V_{sat} = \pm 10$ V, what is the voltage compliance of your circuit?

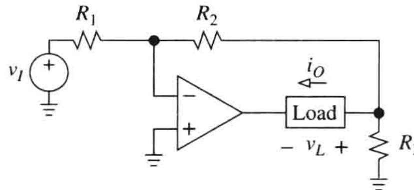


FIGURE P2.7

2.8 (*a*) Except for a finite gain of $a = 10^3$ V/V, the op amp of Fig. P2.7 is ideal. If $R_1 = R_2 = 100$ kΩ, find R_3 for a sensitivity of 1 mA/V. What are the element values of the Norton equivalent seen by the load? (*b*) If $v_I = 2.0$ V, what is the value of i_O for $v_L = 5$ V? For $v_L = -4$ V?

2.9 Consider the following statements about the resistance R_o seen by the load in the V-I converter of Fig. 2.4*b*, where the op amp is assumed ideal: (*a*) Looking toward the left, the load sees $R \parallel r_d = R \parallel \infty = R$, and looking to the right, it sees $r_o = 0$; hence, $R_o = R + 0 = R$. (*b*) Looking toward the left, the load sees a virtual-ground node with zero resistance, and looking to the right, it sees $r_o = 0$; hence, $R_o = 0 + 0 = 0$. (*c*) $R_o = \infty$ because of negative feedback. Which statement is correct? How would you refute the other two?

2.10 (*a*) Assuming ideal op amp in the Howland pump of Fig. 2.6*a*, show that under the condition of Eq. (2.9), the resistance R_i seen by the source v_I can be positive, negative, or even infinite, depending on the load resistance R_L. (*b*) Justify physically for the simplified case $R_1 = R_2 = R_3 = R_4$.

2.11 (*a*) Use Blackman's formula to verify the expression for R_o of Eq. (2.14). (*b*) Suppose a Howland pump is implemented with four 1.0-kΩ matched resistors and an op amp having $a = 10^4$ V/V. If $v_I = -1.0$ V, what is i_O for $v_L = 0$? For $v_L = +5$ V? For $v_L = -2.5$ V?

2.12 Using a 741 powered from ±15-V regulated supplies, design a Howland pump that will *sink* (rather than source) 1.5 mA from a grounded load with a voltage compliance of

± 10 V. To make the circuit self-contained, use the -15-V supply voltage as the input v_I. Then find the current through R_1 and R_2 if the load is (*a*) a 2-kΩ resistor; (*b*) a 6-kΩ resistor; (*c*) a 5-V Zener diode with the cathode at ground; (*d*) a short circuit; (*e*) a 10-kΩ resistor. In (*e*), is i_O still 1.5 mA? Explain.

2.13 Suppose in the Howland circuit of Fig. 2.6*a* we lift the left terminal of R_3 off ground and simultaneously apply an input v_1 via R_3 and an input v_2 via R_1. Show that the circuit is a *difference V-I converter* with $i_O = (1/R_1)(v_2 - v_1) - (1/R_o)v_L$, where R_o is given by Eq. (2.8).

2.14 Design a grounded-load *V-I* converter that converts a 0-V to 10-V input to a 4-mA to 20-mA output. The circuit is to be powered from ± 15-V regulated supplies.

2.15 Design a grounded-load current generator meeting the following specifications: i_O is to be variable over the range $-2\text{ mA} \le i_O \le +2\text{ mA}$ by means of a 100-kΩ pot; the voltage compliance must be 10 V; the circuit is to be powered from ± 15-V regulated supplies.

2.16 (*a*) Prove Eq. (2.15). (*b*) Using a 741 op amp powered from ± 15-V supplies, design an improved Howland circuit with a sensitivity of 1 mA/V for $-10\text{ V} \le v_I \le 10\text{ V}$. Your circuit must work properly over the range $-10\text{ V} \le v_L \le +10\text{ V}$.

2.17 Suppose the improved Howland pump of Fig. 2.11 is implemented with $R_1 = R_3 = R_4 = 20.0$ kΩ, $R_{2A} = R_{2B} = 10.0$ kΩ, and an op amp with $a = 10^4$ V/V. Find R_o via the test method, and check your result via Blackman's formula.

2.18 Suppose the Howland pump of Example 2.4 drives a 0.1-μF load. (*a*) Assuming the capacitance is initially discharged, sketch and label $v_O(t \ge 0)$, and find the time it takes for the op amp to reach saturation. (*b*) Repeat if R_4 is decreased by 10%.

2.19 Sketch and label $v_O(t \ge 0)$ if the Howland pump of Example 2.4 drives an initially discharged 1-μF capacitor, and R_4 is increased by 10%.

2.20 Design an improved Howland circuit whose sensitivity is variable from 0.1 mA/V to 1 mA/V by means of a 100-kΩ pot.

2.21 (*a*) Given that the circuit of Fig. P2.21 yields $i_O = A(v_2 - v_1) - (1/R_o)v_L$, find expressions for A and R_o, as well as the condition among the resistances that yields $R_o = \infty$. (*b*) Discuss the effect of using 1% resistances.

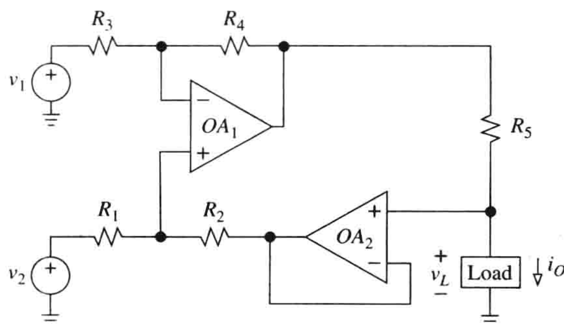


FIGURE P2.21

2.22 (*a*) Given that the circuit of Fig. P2.22 yields $i_O = Av_I - (1/R_o)v_L$, find expressions for A and R_o, as well as the condition among its resistances that yields $R_o = \infty$. (*b*) Discuss the effect of using 1% resistances.

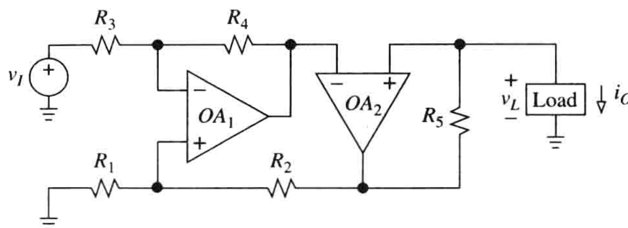


FIGURE P2.22

2.23 Repeat Problem 2.22 for the circuit of Fig. P2.23.

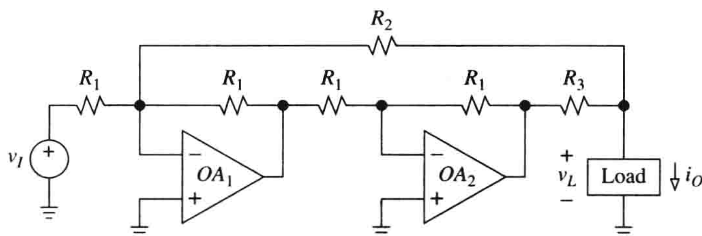


FIGURE P2.23

2.3 Current amplifiers

2.24 (*a*) Prove Eq. (2.18). (*b*) Assuming a 741 op amp in Fig. 2.12*a*, specify resistances for $A = 10$ A/A; estimate the gain error as well as the output resistance of the circuit.

2.25 Find the gain as well as the output impedance of the current amplifier of Fig. P2.25.

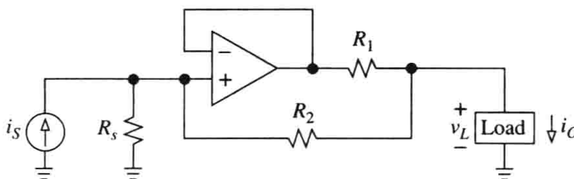


FIGURE P2.25

2.26 Show that if $R_s = \infty$ and $a \neq \infty$ in the current amplifier of Fig. 2.12*b*, then Eq. (2.18) holds.

2.27 A grounded-load current amplifier can be implemented by cascading an I-V and a V-I converter. Using resistances no greater than 1 MΩ, design a current amplifier with $R_i = 0$, $A = 10^5$ A/A, $R_o = \infty$, and a full-scale input of 100 nA. Assuming ±10-V saturation voltages, the voltage compliance must be 5 V.

2.28 Suitably modify the circuit of Fig. P2.23 so that it becomes a current amplifier with $R_i = 0$, $A = 100$ A/A, and $R_o = \infty$. Assume ideal op amps.

2.29 In Fig. P2.29 the odd-numbered inputs are fed to OA_2's summing junction directly, and the even-numbered inputs are fed via a current reverser. Obtain a relationship between v_O and the various inputs. What happens if any of the inputs are left floating? Will it affect the contribution from the other inputs? What is an important advantage of this circuit compared to that of Problem 1.33?

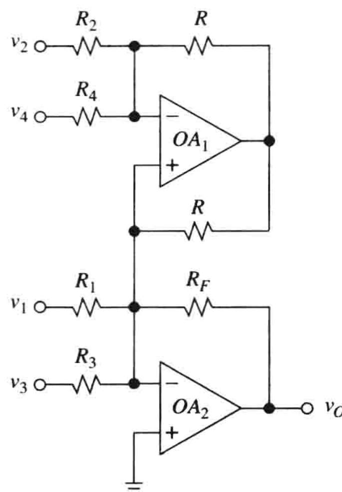


FIGURE P2.29

2.4 Difference amplifiers

2.30 Derive Eq. (2.23).

2.31 (*a*) Derive Eq. (2.27). (*b*) Using a 100-kΩ pot, specify suitable resistances such that varying the wiper from end to end varies the gain from 10 V/V to 100 V/V.

2.32 (*a*) Derive Eq. (2.28). (*b*) Specify suitable component values such that gain can be varied from 1 V/V to 100 V/V.

2.33 (*a*) A difference amplifier has $v_1 = 10\cos(2\pi 60t)$ V $- 5\cos(2\pi 10^3 t)$ mV, and $v_2 = 10\cos(2\pi 60t)$ V$+5\cos(2\pi 10^3 t)$ mV. If $v_O = 100\cos(2\pi 60t)$ mV$+2\cos(2\pi 10^3 t)$ V, find A_{dm}, A_{cm}, and CMRR_{dB}. (*b*) Repeat (*a*) with $v_1 = 10.01\cos(2\pi 60t)$ V $- 5\cos(2\pi 10^3 t)$ mV, $v_2 = 10.00\cos(2\pi 60t)$ V $+ 5\cos(2\pi 10^3 t)$ mV, and $v_O = 0.5\cos(2\pi 60t)$ V $+ 2.5\cos(2\pi 10^3 t)$ V.

2.34 If the actual resistance values in Fig. 2.13*a* are found to be $R_1 = 1.01$ kΩ, $R_2 = 99.7$ kΩ, $R_3 = 0.995$ kΩ, and $R_4 = 102$ kΩ, estimate A_{dm}, A_{cm}, and CMRR_{dB}.

2.35 If the difference amplifier of Fig. 2.13*a* has a differential-mode gain of 60 dB and $\text{CMRR}_{dB} = 100$ dB, find v_O if $v_1 = 4.001$ V and $v_2 = 3.999$ V. What is the percentage error of the output due to finite CMRR?

2.36 If the resistance pairs are perfectly balanced and the op amp is ideal in the difference amplifier of Fig. 2.13*a*, then we have $\text{CMRR}_{dB} = \infty$. But what if the open-loop gain a is finite, everything else being ideal? Is the CMRR still infinite? Justify your finding intuitively.

2.37 A student is trying out the difference amplifier of Fig. 2.13*a* in the lab. The student uses two pairs of perfectly matched resistances, $R_3 = R_1 = 1.0\,\text{k}\Omega$ and $R_4 = R_2 = 100\,\text{k}\Omega$, thus expecting CMRR $= \infty$. (*a*) A peer decides to play a prank by installing an additional 1-MΩ resistor between nodes v_1 and v_P. How is the CMRR affected? (*b*) Repeat, if the 1-MΩ resistor is installed between node v_N and ground. (*c*) Repeat, if the 1-MΩ resistor is installed between nodes v_O and v_P.

2.38 Suppose the circuit of Fig. P2.29 is used with only the v_1 and v_2 inputs, and is implemented with 10-kΩ resistors throughout to give $v_O = v_2 - v_1$ with CMRR $= \infty$. (*a*) Investigate the effect on the CMRR if we use 0.1% resistors and otherwise ideal op amps. (*b*) Investigate the effect of using perfectly matched resistors but op amps with gains of $a = 10^4$ V/V.

2.39 If implemented with $R_1 = R_2 = R_3 = R_4$ and ideal op amps, the circuit of Fig. P2.21 gives $i_O = (v_2 - v_1)/R_5$ regardless of the load. As such, it is a difference amplifier, and therefore prone to CMRR limitations if the resistances are mismatched and op amps have $a \neq \infty$. For the present circuit we define $A_{\text{dm}} = i_{\text{DM(SC)}}/v_{\text{DM}}$ and $A_{\text{cm}} = i_{\text{CM(SC)}}/v_{\text{CM}}$, where $i_{\text{DM(SC)}}$ and $i_{\text{CM(SC)}}$ are, respectively, the short-circuit differential-mode and common-mode output currents, and v_{DM} and v_{CM} are as defined in Eq. (2.21). (*a*) Investigate the effect on the CMRR if we use 1% resistors and otherwise ideal op amps. (*b*) Investigate the effect of using perfectly matched resistors but op amps with gains of $a = 10^3$ V/V.

2.5 Instrumentation amplifiers

2.40 In the IA of Fig. 2.20 let $R_3 = 1\ \text{M}\Omega$, $R_G = 2\ \text{k}\Omega$, and $R_1 = R_2 = 100\ \text{k}\Omega$. If v_{DM} is an ac voltage with a peak amplitude of 10 mV and v_{CM} is a dc voltage of 5 V, find all node voltages in the circuit.

2.41 Show that if OA_1 and OA_2 in Fig. 2.20 have the same open-loop gain a, together they form a negative-feedback system with input $v_I = v_1 - v_2$, output $v_O = v_{O1} - v_{O2}$, open-loop gain a, and feedback factor $\beta = R_G/(R_G + 2R_3)$.

2.42 A triple-op-amp IA is to be implemented with $A = A_I \times A_{II} = 50 \times 20 = 10^3$ V/V. Assuming matched input-stage op amps, find the minimum open-loop gain required of each op amp for a 0.1% maximum deviation of A from the ideal.

2.43 Compared with the classical triple-op-amp IA, the IA of Fig. P2.43 (see *EDN,* Oct. 1, 1992, p. 115) uses fewer resistances. The wiper, nominally positioned halfway, is used to maximize the CMRR. Show that $v_O = (1 + 2R_2/R_1)(v_2 - v_1)$.

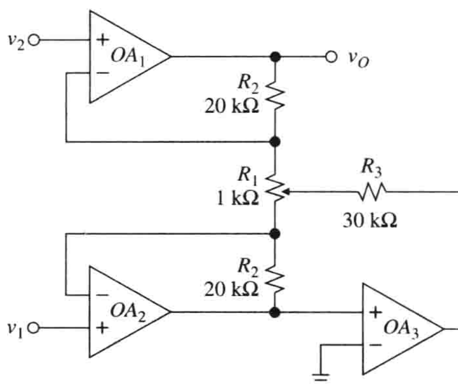


FIGURE P2.43

2.44 (*a*) To investigate the effect of mismatched resistances in the IA of Fig. 2.23, assume that $R_3/R_4 = (R_1/R_2)(1 - \epsilon)$. Show that $v_O = A_{dm}v_{DM} + A_{cm}v_{CM}$, where $A_{dm} = 1 + R_2/R_1 - \epsilon/2$ and $A_{cm} = \epsilon$. (*b*) Discuss the implications of using 1% resistors without trimming for the case $A = 10^2$ V/V.

2.45 (*a*) Derive Eq. (2.35). (*b*) Specify suitable components such that A can be varied over the range $10\ \text{V/V} \le A \le 100\ \text{V/V}$ by means of a 10-kΩ pot.

2.46 The gain of the dual-op-amp IA of Fig. P2.46 (see *EDN,* Feb. 20, 1986, pp. 241–242) is adjustable by means of a single resistor R_G. (*a*) Show that $v_O = 2(1 + R/R_G)(v_2 - v_1)$. (*b*) Specify suitable components to make A variable from 10 V/V to 100 V/V by means of a 10-kΩ pot.

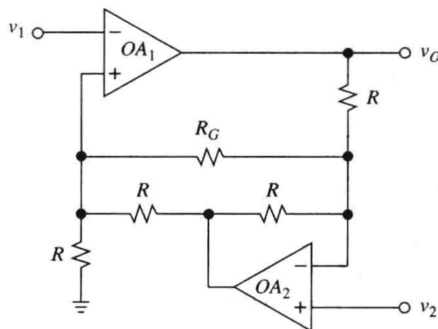


FIGURE P2.46

2.47 The dual-op-amp IA of Fig. P2.47 (see *Signals and Noise, EDN,* May 29, 1986) offers the advantage that by proper adjustment of the pot, a fairly high CMRR can be achieved and maintained well into the kilohertz range. Show that $v_O = (1 + R_2/R_1)(v_2 - v_1)$.

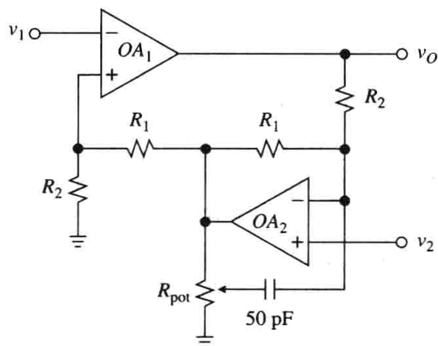


FIGURE P2.47

2.48 Assuming perfectly matched resistances as well as perfectly matched op amps in the dual-op-amp IA of Fig. 2.23, investigate the effect of finite open-loop op amp gain a upon the CMRR of the circuit (except for their finite gain, both op amps are ideal). Assuming $a = 10^5$ V/V, find CMRR_{dB} if $A = 10^3$ V/V. Repeat, but if $A = 10$ V/V, and comment on your findings.

2.49 A technician is assembling the dual-op amp IA of Fig. 2.23 using two pairs of perfectly matched resistances, $R_3 = R_1 = 2.0\ \mathrm{k}\Omega$ and $R_4 = R_2 = 18\ \mathrm{k}\Omega$. (*a*) If he inadvertently installs an additional 1-MΩ resistor between OA_2's inverting input v_{N2} and ground, how is the CMRR affected? (*b*) What if the same resistor is connected between v_{N2} and v_1 instead?

2.6 Instrumentation applications

2.50 Design a digitally programmable IA having an overall gain of 1 V/V, 10 V/V, 100 V/V, and 1000 V/V. Show the final design.

2.51 Assuming ±15-V regulated power supplies, design a programmable IA with two operating modes: in the first mode the gain is 100 V/V and the output offset is 0 V; in the second mode the gain is 200 V/V and the output offset is −5 V.

2.52 (*a*) Derive Eq. (2.42). (*b*) In the current-output IA of Fig. 2.33 specify suitable components for a sensitivity of 1 mA/V. (*c*) Investigate the effect of using 0.1% resistances.

2.53 In the circuit of Fig. 2.33 let $R_1 = R_4 = R_5 = 10\ \mathrm{k}\Omega$, $R_2 = 1\ \mathrm{k}\Omega$, and $R_3 = 9\ \mathrm{k}\Omega$. If an additional resistance R_G is connected between the inverting input nodes of the two op amps, find the gain as a function of R_G.

2.54 (*a*) Design a current-output IA whose sensitivity can be varied from 1 mA/V to 100 mA/V by means of a 100-kΩ pot. The circuit must have a voltage compliance of at least 5 V with ±15-V supplies, and it must have provision for CMRR optimization by means of a suitable trimmer. (*b*) Outline the procedure for calibrating the trimmer.

2.55 Design a current-input, voltage-output IA with a gain of 10 V/mA.

2.7 Transducer bridge amplifiers

2.56 Repeat Example 2.12 using the single-op-amp configuration of Fig. 2.38. Show the final circuit.

2.57 (*a*) Derive Eqs. (2.52) and (2.53). (*b*) Derive Eq. (2.55).

2.58 Assuming that $V_{REF} = 2.5$ V in Fig. 2.39, specify suitable component values for an output sensitivity of 0.1 V/°C with a Pt RTD.

2.59 (*a*) Assuming that $V_{REF} = 15$ V in Fig. 2.40, specify suitable component values for an output sensitivity of 0.1 V/°C with a Pt RTD. (*b*) Assuming the same tolerances as in Example 2.13, make provisions for bridge calibration.

2.60 Show that the linearized bridge circuit of Fig. P2.60 yields $v_O = -RV_{REF}\delta/(R_1 + R)$. Name a disadvantage of this circuit.

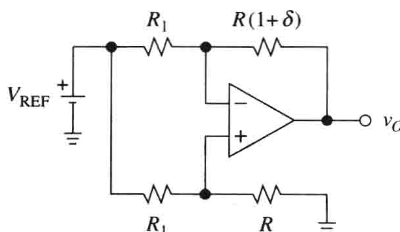


FIGURE P2.60

2.61 Using the circuit of Fig. P2.60 with $V_{REF} = 2.5$ V and an additional gain stage, design an RTD amplifier circuit with a sensitivity of 0.1 V/°C. The circuit is to have provisions for bridge calibration. Outline the calibration procedure.

2.62 Show that the linearized bridge circuit[7] of Fig. P2.62 (U.S. Patent 4,229,692) yields $v_O = R_2 V_{REF}\delta / R_1$. Discuss how you would make provisions for calibrating the circuit.

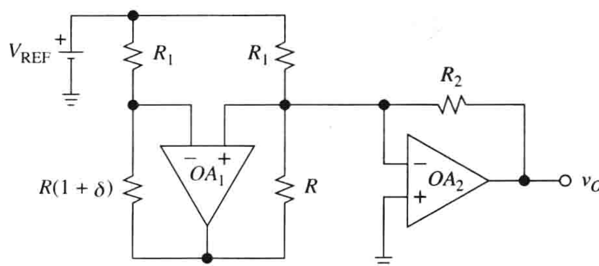


FIGURE P2.62

REFERENCES

1. J. Graeme, "*Photodiode Amplifiers–Op Amp Solutions,* McGraw-Hill, New York, 1996, ISBN-10-007024247X.
2. C. Kitchin and L. Counts, *A Designer's Guide to Instrumentation Amplifiers,* 3d ed., Analog Devices, Norwood, MA, 2006, http://www.analog.com/static/imported-files/design_handbooks/5812756674312778737Complete_In_Amp.pdf.
3. J. Williams, "Applications for a Switched-Capacitor Instrumentation Building Block," Linear Technology Application Note AN-3, http://cds.linear.com/docs/en/application-note/an03f.pdf.
4. N. P. Albaugh, "The Instrumentation Amplifier Handbook, Including Applications," http://www.cypress.com/?docID=38317.
5. Analog Devices Engineering Staff, *Practical Design Techniques for Sensor Signal Conditioning,* Analog Devices, Norwood, MA, 1999, ISBN-0-916550-20-6.
6. "Practical Temperature Measurements," Application Note 290, Hewlett-Packard, Palo Alto, CA, 1980.
7. J. Graeme, "Tame Transducer Bridge Errors with Op Amp Feedback Control," *EDN,* May 26, 1982, pp. 173–176.
8. J. Williams, "Good Bridge-Circuit Design Satisfies Gain and Balance Criteria," *EDN,* Oct. 25, 1990, pp. 161–174.

3

ACTIVE FILTERS: PART I

滤波器是一种在频率基础上处理信号的电路。它根据频率变化而变化的行为特性称为频率响应，并以传递函数 $H(j\omega)$ 表示，这里 $\omega=2\pi f$ 是以弧度 / 秒为单位的角频率，j 是虚数单位（$j^2=-1$）。频率响应可以进一步具体分为幅度响应 $|H(j\omega)|$ 和相位响应 $\measuredangle H(j\omega)$；它们分别给出了当交流信号通过滤波器所产生的幅度增益和相位偏移。

频率响应概述

根据响应幅度，滤波器可以分为低通、高通、带通和带阻（或者陷波）滤波器。第 5 类滤波器是只处理相位而幅度不变的全通滤波器。参考图 3.1，我们将按照以下方式定义这些理想响应。

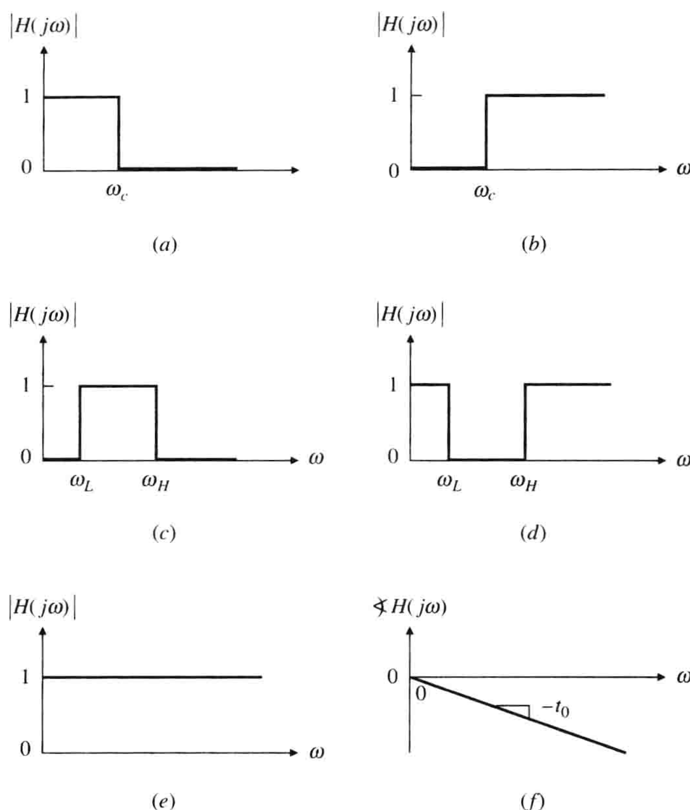


图 3.1 理想的滤波器响应：(a) 低通，(b) 高通，(c) 带通，(d) 带阻，以及 (e)、(f) 全通滤波器

低通响应通过一个称为截止频率 ω_C 的频率来表征。当频率 $\omega<\omega_C$ 时，$|H|=1$；当频率 $\omega>\omega_C$ 时，$|H|=0$。这表明频率小于 ω_C 的输入信号通过滤波器后幅度没有变化，而频率大于 ω_C 的输入信号通过滤波器后完全衰减。低通滤波器通常用于去除一个信号中的高频噪声。

高通响应正好与低通响应相反。频率大于 ω_C 的输入信号通过滤波器后幅度没有变化，而频率小于 ω_C 的输入信号被滤波器完全阻挡。

带通响应通过一个称为通带的频率带宽 $\omega_L<\omega<\omega_H$ 来表征。频率在通带之内的输入信号通过滤波器后幅度没有变化；在通带之外（$\omega<\omega_L$ 或者 $\omega>\omega_H$）的输入信号被完全截止。一种熟知的带通滤波器就是收音机的调谐电路，它可以让用户选择某一个特定的电台而屏蔽掉其他的电台。

带阻响应正好与带通响应相反。因为它阻挡了频率在阻带 $\omega_L<\omega<\omega_H$ 范围内的输入信号，而完全通过其他频率（$\omega<\omega_L$ 或者 $\omega>\omega_H$）的信号。当阻带足够窄的时候，带阻响应也称为陷波响应。陷波滤波器的一种应用就是去除医疗设备上拾取到的不需要的 60Hz 频率干扰分量。

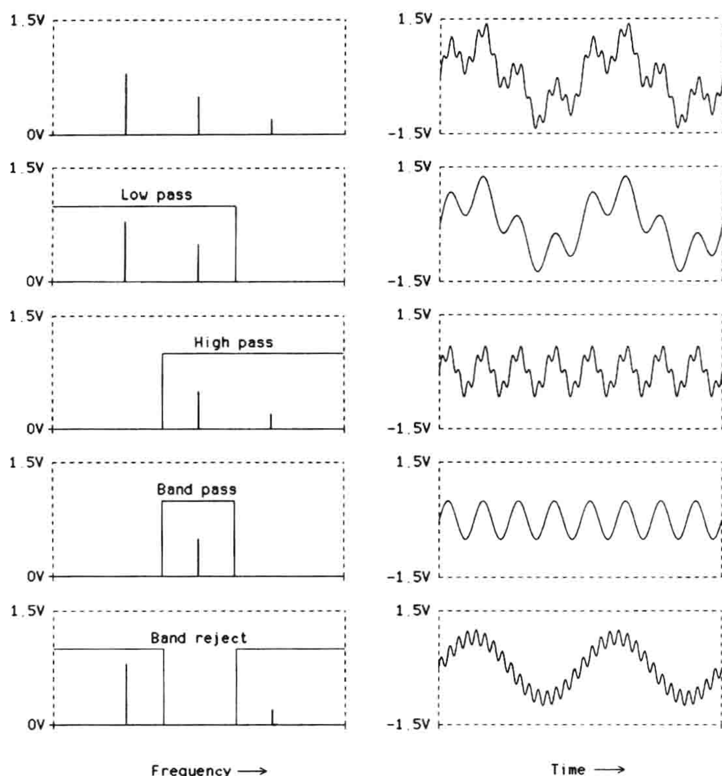


图 3.2　频率域（左边）和时域（右边）的滤波效果

全通响应由与频率无关的幅度增益 $|H|=1$ 和相移 $\measuredangle H=-t_0\omega$ 来表征，其中 t_0 是一个以秒为单位的合适比例常数。这种滤波器对通过的交流信号的幅值完全没有影响，但是根据频率 ω 对信号进行延时。显而易见，全通滤波器也可以称为延时滤波器。延时均衡器和宽带 90° 相移网络都是全通滤波器的例子。

作为一个例子，图 3.2 说明了前 4 种理想滤波器对于输入电压

$$v_I(t) = 0.8\sin\omega_0 t + 0.5\sin\omega_0 t + 0.2\sin 16\omega_0 t \text{ V}$$

的滤波效果。左图所示的是由频谱分析仪观测到的信号频谱，右图所示的是由示波器观测到的信号时域波形。其中，上面是输入信号的频谱和波形，而下面分别是低通、高通、带通和带阻滤波器的输出信号的频谱和波形。比如，如果输入信号 $v_I(t)$ 通过一个截止频率 ω_C 为 $4\omega_0$ ~ $16\omega_0$ 之间的某个值的低通滤波器，那么输入信号的前两个分量的幅度乘以 1，但是第三个分量乘以 0 而被阻挡，结果输出电压就是 $v_O(t) = 0.8\sin\omega_0 t + 0.5\sin 4\omega_0 t$ V。

随着继续讨论，我们会发现实际的滤波器仅仅是图 3.2 中所示的理想砖墙幅度变化滤波器的近似，而且它们同时会引起相位变化。

有源滤波器

滤波器理论是一门庞大的课程，有多本专门研究它的教科书。滤波器可以单单通过电阻、电感、电容等无源器件组成（*RLC* 滤波器）。然而，在反馈的概念出现之后，滤波器可以通过使用放大器而不使用电感实现几乎所有的响应。因为在基本的电路元件中，电感是最不理想的，并且电感体积大、重量重、成本高，不适合用于 IC 的大规模量产，所以这是一个很大的优势。

如何处理运算放大器来取代电感是接下来要讨论的一个令人感兴趣的问题。因为运算放大器可以从电源吸收能量，然后将能量释放到周围的电路来弥补电阻消耗的能量，所以我们可以直观地认为这是合理的。电感和电容是不损耗能量的元件，它们在一个周期的一部分时间内吸收能量，在另一部分时间内释放能量。由电源支持的运算放大器能做同样的或者更多的事情。与电容或者电感不同的是，运算放大器可以释放比电阻实际消耗更多的能量。因此，运算放大器称为*有源器件*，采用放大器组成的滤波器称为*有源滤波器*。这些滤波器为运算放大器提供了一个最为丰富的应用平台。

有源滤波器只能够在运算放大器能够正常工作的范围之内工作。运算放大器最严重的限制是随着频率的增加，开环增益衰减。这将在第 6 章详细讨论。这一般将有源滤波器的应用范围限制在 100MHz 以下。这包括了音频和仪器仪用的应用范围。那是有源滤波器最为广泛应用的范围，而电感由于过于笨重而无法与小型化的 IC 进行竞争。在运算放大器的工作范围之外，电感又将占优势，所以高频滤波器还使用无源 *RLC* 元件构成。在这些滤波器中，由于电感值和电容值随着工作频率范围的增高而减小，电感的尺寸和重量将变得更加便于处理。

本章重点

本章将研究一阶和二阶有源滤波器（高阶滤波器和开关电容滤波器将在第 4 章研究）。首先介绍一般滤波器概念，之后将进入一阶滤波器及其典型应用的介绍，比如音频应用。接着进入现在日常使用的二阶有源滤波器介绍，比如 *KRC* 滤波器、多重反馈滤波器和状态变量 / 双二阶滤波器。本章最后介绍滤波器的灵敏度。本章大量运用了 SPICE 来描绘滤波器的响应，并且就像实验时用示波器一样进行测量演示。

滤波器的主题十分庞大（有好多整本只用来研究滤波器内容的书籍）。所以，这里有必要将内容限制在最受欢迎的滤波器类型上。此外，至少对于初学者而言，有源滤波器的学习是令人不知所措的。所以，我们将故意假设运算放大器是理想的，以便于简化问题的处理。一般感觉而言，只有在掌握了滤波器的概念之后，我们才能够将视线扩张到运算放大器的非理想特性上。这个问题参见第 5 章和第 6 章。

3.1 THE TRANSFER FUNCTION

Filters are implemented with devices exhibiting frequency-dependent characteristics, such as capacitors and inductors. When subjected to ac signals, these elements oppose current flow in a frequency-dependent manner and also introduce a 90° phase shift between voltage and current. To account for this behavior, we use the *complex impedances* $Z_L = sL$ and $Z_C = 1/sC$, where $s = \sigma + j\omega$ is the *complex frequency,* in complex nepers per second (complex Np/s). Here, σ is the *neper frequency,* in nepers per second (Np/s) and ω is the *angular frequency,* in radians per second (rad/s).

The behavior of a circuit is uniquely characterized by its transfer function $H(s)$. To find this function, we first derive an expression for the output X_o in terms of the input X_i (X_o and X_i can be voltages or currents) using familiar tools such as Ohm's law $V = Z(s)I$, KVL, KCL, the voltage and current divider formulas, and the superposition principle. Then, we solve for the ratio

$$H(s) = \frac{X_o}{X_i} \tag{3.1}$$

Once $H(s)$ is known, the response $x_o(t)$ to a given input $x_i(t)$ can be found as

$$x_o(t) = \mathcal{L}^{-1}\{H(s)X_i(s)\} \tag{3.2}$$

where $\mathcal{L}^{-1}$ denotes the inverse Laplace transform, and $X_i(s)$ is the Laplace transform of $x_i(t)$.

Transfer functions turn out to be *rational functions* of s,

$$H(s) = \frac{N(s)}{D(s)} = \frac{a_m s^m + a_{m-1}s^{m-1} + \cdots + a_1 s + a_0}{b_n s^n + b_{n-1}s^{n-1} + \cdots + b_1 s + b_0} \tag{3.3}$$

where $N(s)$ and $D(s)$ are suitable polynomials of s with real coefficients and with degrees m and n. The degree of the denominator determines the *order* of the filter (first-order, second-order, etc.). The roots of the equations $N(s) = 0$ and $D(s) = 0$ are called, respectively, the *zeros* and the *poles* of $H(s)$, and are denoted as $z_1, z_2, \ldots, z_m$, and $p_1, p_2, \ldots, p_n$. Factoring out $N(s)$ and $D(s)$ in terms of their respective roots, we can write

$$H(s) = H_0 \frac{(s - z_1)(s - z_2)\cdots(s - z_m)}{(s - p_1)(s - p_2)\cdots(s - p_n)} \tag{3.4}$$

where $H_0 = a_m/b_n$ is called the *scaling factor.* Aside from H_0, $H(s)$ is uniquely determined once its zeros and poles are known. Roots are also referred to as *critical* or *characteristic frequencies* because they depend solely on the circuit, that is, on its elements and the way they are interconnected, irrespective of its signals or the energy stored in its reactive elements. In fact, essential circuit specifications are often given in terms of the roots.

Roots can be real or complex. When zeros or poles are complex, they occur in conjugate pairs. For instance, if $p_k = \sigma_k + j\omega_k$ is a pole, then $p_k^* = \sigma_k - j\omega_k$ is also a pole. Roots are conveniently visualized as points in the *complex plane,* or *s plane:* σ_k is plotted against the horizontal, or *real,* axis, which is calibrated in nepers per second (Np/s); ω_k is plotted against the vertical, or *imaginary,* axis, which is calibrated in

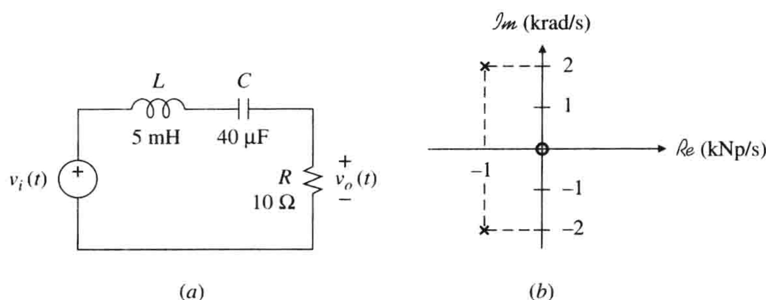


FIGURE 3.3
Circuit of Example 3.1 and its pole-zero plot.

radians per second (rad/s). In these plots a zero is represented as "o" and a pole as "×". Just by looking at the pole-zero pattern of a circuit, a designer can predict important characteristics, such as stability and frequency response. Because these characteristics will arise frequently as we proceed, we wish to give them a definitive review.

EXAMPLE 3.1. Find the pole-zero plot of the circuit of Fig. 3.3*a*.

Solution. Using the generalized voltage divider formula, $V_o = [R/(sL+1/sC+R)]V_i$. Rearranging,

$$H(s) = \frac{V_o}{V_i} = \frac{RCs}{LCs^2 + RCs + 1} = \frac{R}{L} \times \frac{s}{s^2 + (R/L)s + 1/LC}$$

Substituting the given component values and factoring out,

$$H(s) = 2 \times 10^3 \times \frac{s}{[s - (-1 + j2)10^3] \times [s - (-1 - j2)10^3]}$$

This function has $H_0 = 2 \times 10^3$ V/V, a zero at the origin, and a conjugate pole pair at $-1 \pm j2$ complex kNp/s. Its pole-zero plot is shown in Fig 3.3*b*.

$H(s)$ and Stability

A circuit is said to be stable if it produces a bounded output in response to any bounded input. One way to assess whether a circuit is stable or not is to inject some energy into one or more of its reactive elements and then observe how the circuit does on its own, in the absence of any applied sources. The circuit response is in this case called the *source-free,* or *natural, response.* A convenient method of injecting energy is to apply an impulsive input, whose Laplace transform is unity. By Eq. (3.2), the ensuing response, or impulse response, is then $h(t) = \mathcal{L}^{-1}\{H(s)\}$. Interestingly enough, this response is determined by the poles. We identify two representative cases:

1. $H(s)$ has a real pole at $s = \sigma_k \pm j0 = \sigma_k$. Using well-known Laplace-transform techniques,[5] one can prove that $H(s)$ contains the term $A_k/(s - \sigma_k)$, where A_k is called the *residue* of $H(s)$ at that pole, and is found as $A_k = (s - \sigma_k)H(s)|_{s=\sigma_k}$.

From the Laplace-transform tables we find

$$\mathcal{L}^{-1}\left\{\frac{A_k}{s-\sigma_k}\right\} = A_k e^{\sigma_k t} u(t) \tag{3.5}$$

where $u(t)$ is the unit step function ($u = 0$ for $t < 0$, $u = 1$ for $t > 0$). A real pole contributes an exponential component to the response $x_o(t)$, and this component decays if $\sigma_k < 0$, remains constant if $\sigma_k = 0$, and diverges if $\sigma_k > 0$.

2. $H(s)$ has a complex pole pair at $s = \sigma_k \pm j\omega_k$. In this case $H(s)$ contains the complex term $A_k/[s - (\sigma_k + j\omega_k)]$ as well as its conjugate, and the residue is found as $A_k = [s - (\sigma_k + j\omega_k)]H(s)|_{s=\sigma_k+j\omega_k}$. The inverse Laplace transform of their combination is

$$\mathcal{L}^{-1}\left\{\frac{A_k}{s-(\sigma_k+j\omega_k)} + \frac{A_k^*}{s-(\sigma_k-j\omega_k)}\right\} = 2|A_k|e^{\sigma_k t}u(t)\cos(\omega_k t + \measuredangle A_k) \tag{3.6}$$

This component represents a damped sinusoid if $\sigma_k < 0$, a constant-amplitude, or sustained, sinusoid if $\sigma_k = 0$, and a growing sinusoid if $\sigma_k > 0$.

It is apparent that for a circuit to be stable, *all poles must lie in the left half of the s plane* (LHP), where $\sigma < 0$. Passive *RLC* circuits, such as that of Example 3.1, meet this constraint and are thus stable. However, if a circuit contains dependent sources such as op amps, its poles may spill into the right half-plane and thus lead to instability. Its output will grow until the saturation limits of the op amp are reached. If the circuit has a complex pole pair, the outcome of this is a sustained oscillation. Instability is generally undesirable, and stabilization techniques are covered in Chapter 8. There are nevertheless situations in which instability is exploited on purpose. A common example is the design of sine wave oscillators, to be addressed in Chapter 10.

EXAMPLE 3.2. Find the impulse response of the circuit of Example 3.1.

Solution. We have $A_1 = [s - (-1 + j2)10^3]H(s)|_{s=(-1+j2)10^3} = 1000 + j500 = 500\sqrt{5}\underline{/26.57^\circ}$. So, $v_o(t) = 10^3\sqrt{5}e^{-10^3 t}u(t)\cos(2 \times 10^3 t + 26.57^\circ)$ V.

$H(s)$ and the Frequency Response

In the study of filters we are interested in the response to an ac input of the type

$$x_i(t) = X_{im}\cos(\omega t + \theta_i)$$

where X_{im} is the amplitude, ω the angular frequency, and θ_i the phase angle. In general, the complete response $x_o(t)$ of Eq. (3.2) consists of two components,[5] namely, a *transient* component functionally similar to the natural response, and a *steady-state* component having the same frequency as the input, but differing in amplitude and phase. If all poles are in the LHP, the transient component will die out, leaving only the steady-state component,

$$x_o(t) = X_{om}\cos(\omega t + \theta_o)$$

This is illustrated in Fig. 3.4. Since we are narrowing our scope to this component alone, we wonder whether we can simplify our math, bypassing the general Laplace

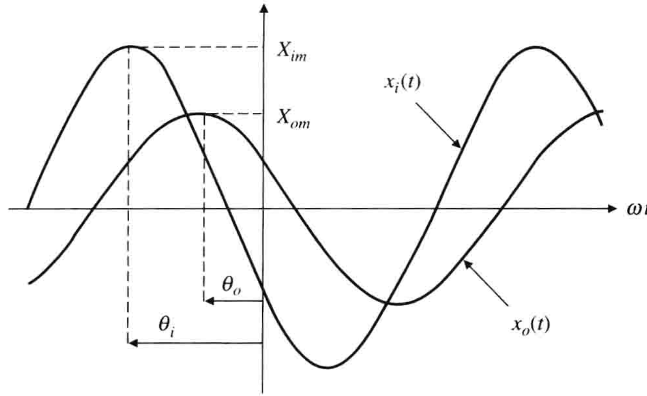


FIGURE 3.4
In general, a filter affects both amplitude and phase.

approach of Eq. (3.2). Such a simplification is possible, and it merely requires that we compute $H(s)$ on the imaginary axis. We do this by letting $s \rightarrow j\omega$ (or $s \rightarrow j2\pi f$ when working with the cyclical frequency *f*, in hertz). Then, the output parameters are found as

$$X_{om} = |H(j\omega)| \times X_{im} \tag{3.7a}$$

$$\theta_o = \sphericalangle H(j\omega) + \theta_i \tag{3.7b}$$

In the course of complex-number manipulations, we shall often use the following important properties: Let

$$H = |H| \underline{/\sphericalangle H} = H_r + jH_i \tag{3.8}$$

where $|H|$ is the *modulus* or *magnitude of H*, $\sphericalangle H$ its *argument* or *phase angle,* and H_r and H_i the *real* and *imaginary* parts. Then,

$$|H| = \sqrt{H_r^2 + H_i^2} \tag{3.9a}$$

$$\sphericalangle H = \tan^{-1}(H_i/H_r) \qquad \text{if } H_r > 0 \tag{3.9b}$$

$$\sphericalangle H = 180^\circ - \tan^{-1}(H_i/H_r) \qquad \text{if } H_r < 0 \tag{3.9c}$$

$$|H_1 \times H_2| = |H_1| \times |H_2| \tag{3.10a}$$

$$\sphericalangle(H_1 \times H_2) = \sphericalangle H_1 + \sphericalangle H_2 \tag{3.10b}$$

$$|H_1/H_2| = |H_1|/|H_2| \tag{3.11a}$$

$$\sphericalangle(H_1/H_2) = \sphericalangle H_1 - \sphericalangle H_2 \tag{3.11b}$$

EXAMPLE 3.3. Find the steady-state response of the circuit of Example 3.1 to the signal $v_i(t) = 10\cos(10^3 t + 45^\circ)$ V.

Solution. Letting $s \to j10^3$ rad/s in Example 3.1 we get $H(j10^3) = j1/(2 + j1) = (1/\sqrt{5})/\underline{/63.43^\circ}$ V/V. So $V_{om} = 10/\sqrt{5}$ V, $\theta_o = 63.43^\circ + 45^\circ = 108.43^\circ$, and $v_o(t) = \sqrt{20}\cos(10^3 t + 108.43^\circ)$ V.

There are various viewpoints we can take in regard to $H(j\omega)$. Presented with the circuit diagram of a filter, we may wish to find $H(s)$ analytically, and then plot $|H(j\omega)|$ and $\measuredangle H(j\omega)$ versus ω (or f) for a visual display of the frequency response. These plots, referred to as *Bode plots,* can be generated by hand or via PSpice.

Conversely, given $H(j\omega)$, we may want to let $j\omega \to s$ to obtain $H(s)$, find its roots, and construct the pole-zero plot.

Alternatively, $H(j\omega)$ may be given to us, either analytically or in graphical form or in terms of filter specifications, and we may be asked to design a circuit realizing this function. The idealized brick-wall responses of Fig. 3.1 cannot be achieved in practice but can be approximated via rational functions of s. The degree n of $D(s)$ determines the order of the filter (first-order, second-order, etc.). As a general rule, the higher n, the greater the flexibility in the choice of the polynomial coefficients best suited to a given frequency-response profile. However, circuit complexity increases with n, indicating a trade-off between how close to ideal we want to be and the price we are willing to pay.

Yet another viewpoint is one in which a filter is given to us in black-box form and we are asked to find $H(j\omega)$ experimentally. By Eq. (3.7), the magnitude and phase are $|H(j\omega)| = X_{om}/X_{im}$ and $\measuredangle H(j\omega) = \theta_o - \theta_i$. To find $H(j\omega)$ experimentally, we apply an ac input and measure the amplitude and phase of the output relative to the input at different frequencies. We then plot measured data versus frequency point-by-point and obtain the experimental profiles of $|H(j\omega)|$ and $\measuredangle H(j\omega)$. If desired, measured data can be processed with suitable curve-fitting algorithms to obtain an analytical expression for $H(j\omega)$ in terms of its critical frequencies. In the case of voltage signals, the measurements are easily done with a dual-trace oscilloscope. To simplify the calculations, it is convenient to set $V_{im} = 1$ V and to adjust the trigger so that $\theta_i = 0$. Then we have $|H(j\omega)| = V_{om}$ and $\measuredangle H(j\omega) = \theta_o$.

Bode Plots

The magnitude and frequency range of a filter can be quite wide. For instance, in audio filters the frequency range is typically from 20 Hz to 20 kHz, which represents a 1000:1 range. In order to visualize small as well as large details with the same degree of clarity, $|H|$ and $\measuredangle H$ are plotted on *logarithmic* and *semilogarithmic* scales, respectively. That is, frequency intervals are expressed in *decades* $(\ldots, 0.01, 0.1, 1, 10, 100, \ldots)$ or in *octaves* $(\ldots, \frac{1}{8}, \frac{1}{4}, \frac{1}{2}, 1, 2, 4, 8, \ldots)$, and $|H|$ is expressed in *decibels* (dB) as

$$|H|_{\text{dB}} = 20\log_{10}|H| \tag{3.12}$$

The Bode plots are plots of decibels and degrees versus decades (or octaves). Another advantage of these plots is that the following useful properties hold:

$$|H_1 \times H_2|_{\rm dB} = |H_1|_{\rm dB} + |H_2|_{\rm dB} \tag{3.13a}$$

$$|H_1/H_2|_{\rm dB} = |H_1|_{\rm dB} - |H_2|_{\rm dB} \tag{3.13b}$$

$$|1/H|_{\rm dB} = -|H|_{\rm dB} \tag{3.13c}$$

To speed up the hand generation of these plots, it is often convenient to effect asymptotic approximations. To this end, the following properties are useful:

$$H \cong H_r \qquad \text{if } |H_r| \gg |H_i| \tag{3.14a}$$

$$H \cong jH_i \qquad \text{if } |H_i| \gg |H_r| \tag{3.14b}$$

Keep Eqs. (3.13) and (3.14) in mind because we shall use them frequently.

3.2 FIRST-ORDER ACTIVE FILTERS

The simplest active filters are obtained from the basic op amp configurations by using a capacitance as one of its external components. Since $Z_C = 1/sC = 1/j\omega C$, the result is a gain with frequency-dependent magnitude and phase. As you study filters, it is important that you try justifying your mathematical findings using physical insight. In this respect, a most valuable tool is asymptotic verification, which is based on the following properties:

$$\lim_{\omega \to 0} Z_C = \infty \tag{3.15a}$$

$$\lim_{\omega \to \infty} Z_C = 0 \tag{3.15b}$$

In words, at low frequencies a capacitance tends to behave as an open circuit compared with the surrounding elements, and at high frequencies it tends to behave as a short circuit.

The Differentiator

In the inverting configuration of Fig. 3.5a we have $V_o = (-R/Z_C)V_i = -RCsV_i$. By a well-known Laplace-transform property, multiplication by s in the frequency domain is equivalent to differentiation in the time domain. This confirms the designation *differentiator* for the circuit. Solving for the ratio V_o/V_i gives

$$H(s) = -RCs \tag{3.16}$$

indicating a zero at the origin.

Letting $s \to j\omega$ and introducing the scaling frequency

$$\omega_0 = \frac{1}{RC} \tag{3.17}$$

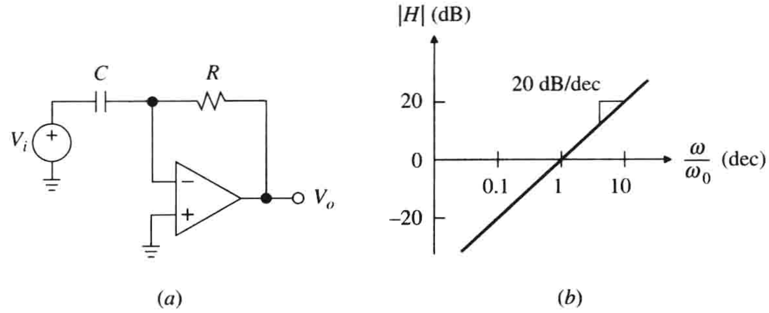


FIGURE 3.5
The differentiator and its magnitude Bode plot.

we can express $H(j\omega)$ in the normalized form

$$H(j\omega) = -j\omega/\omega_0 = (\omega/\omega_0)\,\underline{/-90^\circ} \tag{3.18}$$

Considering that $|H|_{\text{dB}} = 20\log_{10}(\omega/\omega_0)$, the plot of $|H|_{\text{dB}}$ versus $\log_{10}(\omega/\omega_0)$ is a straight line of the type $y = 20x$. As shown in Fig. 3.5*b*, its slope is 20 dB/dec, indicating that for every decade increase (or decrease) in frequency, magnitude increases (or decreases) by 20 dB. Equation (3.18) indicates that the circuit introduces a 90° phase lag and amplifies in proportion to frequency. Physically, we observe that at low frequencies, where $|Z_C| > R$, the circuit provides attenuation (negative decibels); at high frequencies, where $|Z_C| < R$, it provides magnification (positive decibels); at $\omega = \omega_0$, where $|Z_C| = R$, it provides unity gain (0 db). Consequently, ω_0 is called the *unity-gain frequency*.

Integrators

Also called a *Miller integrator* because the capacitor is in the feedback path, the circuit of Fig. 3.6*a* gives $V_o = (-Z_C/R)V_i = -(1/RCs)V_i$. The fact that division by s in the frequency domain corresponds to integration in the time domain confirms the designation *integrator*. Its transfer function

$$H(s) = -\frac{1}{RCs} \tag{3.19}$$

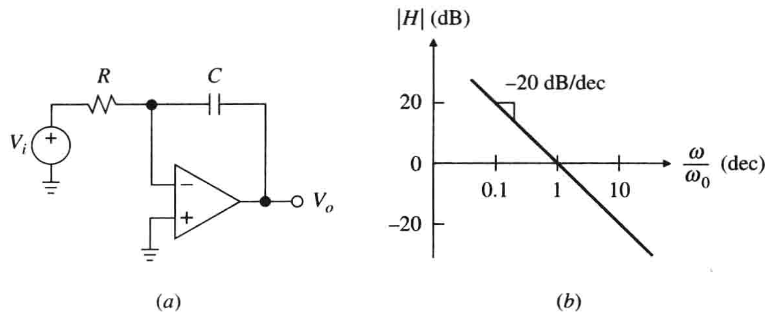


FIGURE 3.6
The integrator and its magnitude Bode plot.

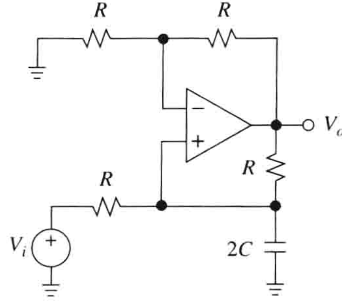


FIGURE 3.7
Noninverting, or Deboo, integrator.

has a pole at the origin. Letting $s \to j\omega$, we can write

$$H(j\omega) = -\frac{1}{j\omega/\omega_0} = \frac{1}{\omega/\omega_0}\angle{+90^\circ} \tag{3.20}$$

where $\omega_0 = 1/RC$, as in Eq. (3.17). Observing that the transfer function is the reciprocal of that of the differentiator, we can apply Eq. (3.13*c*) and construct the integrator magnitude plot simply by reflecting that of the differentiator about the 0-dB axis. The result, shown in Fig. 3.6*b,* is a straight line with a slope of −20 dB/dec and with ω_0 as the *unity-gain frequency.* Moreover, the circuit introduces a 90° phase lead.

Because of the extremely high gain at low frequencies, where $|Z_C| \gg R$, a practical integrator circuit is seldom used alone as it tends to saturate. As mentioned in Chapter 1, an integrator is usually placed inside a control loop designed to keep the op amp within the linear region. We shall see examples when studying state-variable and biquad filters in Section 3.7 and sine wave oscillators in Section 10.1.

Because of the negative sign in Eq. (3.19), the Miller integrator is also said to be an *inverting integrator.* The circuit of Fig. 3.7, called the *Deboo integrator,* for its inventor, uses a Howland current pump with a capacitance as load to achieve noninverting integration. As we know, the pump forces a current $I = V_i/R$ into the capacitance, resulting in a noninverting-input voltage $V_p = (1/s2C)I = V_i/2sRC$. The op amp then amplifies this voltage to give $V_o = (1 + R/R)V_p = V_i/sRC$, so

$$H(s) = \frac{1}{RCs} \tag{3.21}$$

The magnitude plot is the same as for the inverting integrator. However, the phase angle is now −90°, rather than +90°.

It is instructive to investigate the circuit from the more general viewpoint of Fig. 3.8*a*, where we identify two blocks: the *RC* network shown at the bottom and the rest of the circuit forming a negative resistance converter. The converter provides a variable resistance $-R(R/kR) = -R/k$, $k \geq 0$, so the net resistance seen by C is $R \parallel (-R/k) = R/(1-k)$, indicating the pole

$$p = -\frac{1-k}{RC} \tag{3.22}$$

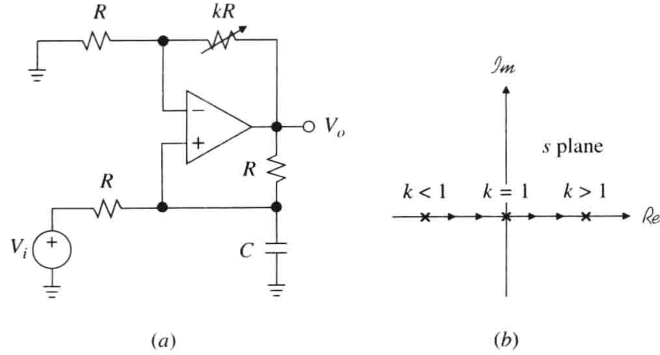


FIGURE 3.8
Varying k varies pole location.

The natural response is then

$$v_O(t) = v_O(0)e^{-t(1-k)/RC}u(t) \tag{3.23}$$

We identify three important cases: (a) For $k < 1$, positive resistance prevails, indicating a negative pole and an exponentially decaying response. The decay is due to dissipation of the energy stored in the capacitance by the net resistance. (b) For $k = 1$, the energy supplied by the negative resistance balances the energy dissipated by the positive resistance, yielding a constant response. The net resistance is now infinite, and the pole is right at the origin. (c) For $k > 1$, the negative resistance supplies more energy than the positive resistance can dissipate, causing an exponential buildup. Negative resistance prevails, the pole is now in the right half plane, and the response diverges. Figure 3.8b shows the root locus as k is increased.

Low-Pass Filter with Gain

Placing a resistor in parallel with the feedback capacitor, as in Fig. 3.9a, turns the integrator into a low-pass filter with gain. Letting $1/Z_2 = 1/R_2 + 1/(1/sC) = (R_2Cs + 1)/R_2$ gives $H(s) = -Z_2/R_1$, or

$$H(s) = -\frac{R_2}{R_1}\frac{1}{R_2Cs + 1} \tag{3.24}$$

indicating a real pole at $s = -1/R_2C$. Letting $s \to j\omega$, we can express $H(s)$ in the normalized form

$$H(j\omega) = H_0\frac{1}{1 + j\omega/\omega_0} \tag{3.25a}$$

$$H_0 = -\frac{R_2}{R_1} \qquad \omega_0 = \frac{1}{R_2C} \tag{3.25b}$$

Physically, the circuit works as follows. At sufficiently low frequencies, where $|Z_C| \gg R_2$, we can ignore Z_C compared with R_2 and thus regard the circuit as an inverting amplifier with gain $H \cong -R_2/R_1 = H_0$. For obvious reasons, H_0 is called the *dc gain.* As shown in Fig. 3.9b, the low-frequency asymptote of the magnitude Bode plot is a horizontal line positioned at $|H_0|_{\text{dB}}$.

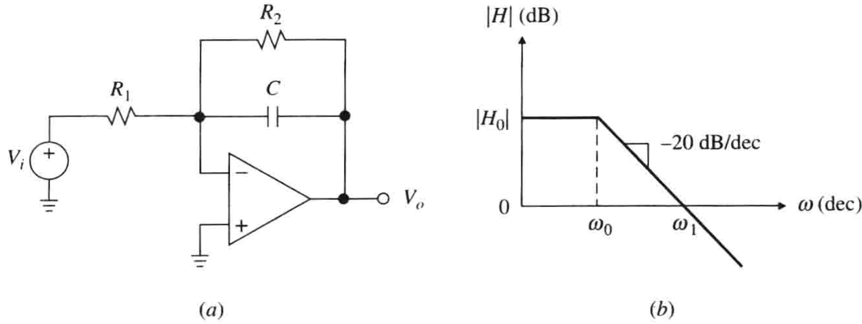


FIGURE 3.9
Low-pass filter with gain.

At sufficiently high frequencies, where $|Z_C| \ll R_2$, we can ignore R_2 compared with Z_C and thus regard the circuit as an integrator. As we know, its high-frequency asymptote is a line with a slope of −20 dB/dec and passing through the unity-gain frequency $\omega_1 = 1/R_1C$. Since the circuit approximates integrator behavior over only a limited frequency range, it is also called a *lossy integrator.*

The borderline between amplifier and integrator behavior occurs at the frequency that makes $|Z_C| = R_2$, or $1/\omega C = R_2$. Clearly, this is the frequency ω_0 of Eq. (3.25*b*). For $\omega/\omega_0 = 1$, Eq. (3.25*a*) predicts $|H| = |H_0/(1 + j1)| = |H_0|/\sqrt{2}$, or, equivalently, $|H|_{\text{dB}} = |H_0|_{\text{dB}} - 3$ dB. Hence, ω_0 is called the *−3-dB frequency.*

The magnitude profile indicates that this is a low-pass filter with H_0 as dc gain and with ω_0 as cutoff frequency. Signals with $\omega < \omega_0$ are passed with gain close to H_0, but signals with $\omega > \omega_0$ are progressively attenuated, or cut. For every decade increase in ω, $|H|$ decreases by 20 dB. Clearly, this is only a crude approximation to the brick-wall profile of Fig. 3.1*b*.

EXAMPLE 3.4. (*a*) In the circuit of Fig. 3.9*a*, specify suitable components to achieve a −3-dB frequency of 1 kHz with a dc gain of 20 dB and an input resistance of at least 10 kΩ. (*b*) At what frequency does gain drop to 0 dB? What is the phase there?

Solution.

(*a*) Since 20 dB corresponds to $10^{20/20} = 10$ V/V, we need $R_2 = 10R_1$. To ensure $R_i > 10$ kΩ, try $R_1 = 20$ kΩ. Then, $R_2 = 200$ kΩ, and $C = 1/\omega_0 R_2 = 1/(2\pi \times 10^3 \times 200 \times 10^3) = 0.796$ nF. Use $C = 1$ nF, which is a more readily available value. Then, scale the resistances as $R_2 = 200 \times 0.796 = 158$ kΩ and $R_1 = 15.8$ kΩ, both 1%.

(*b*) Imposing $|H| = 10/\sqrt{1^2 + (f/10^3)^2} = 1$ and solving yields $f = 10^3\sqrt{10^2 - 1} = 9.950$ kHz. Moreover, $\measuredangle H = 180° - \tan^{-1} 9950/10^3 = 95.7°$.

High-Pass Filter with Gain

Placing a capacitor in series with the input resistor as in Fig. 3.10*a* turns the differentiator into a high-pass filter with gain. Letting $Z_1 = R_1 + 1/sC = (R_1Cs + 1)/sC$ and $H(s) = -R_2/Z_1$ gives

$$H(s) = -\frac{R_2}{R_1}\frac{R_1Cs}{R_1Cs + 1} \tag{3.26}$$

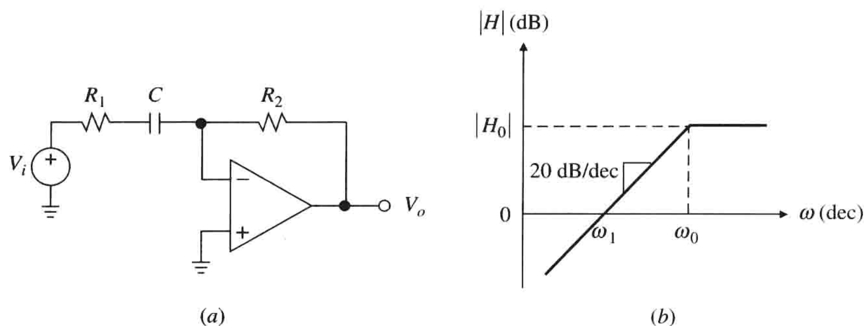


FIGURE 3.10
High-pass filter with gain.

indicating a zero at the origin and a real pole at $s = -1/R_1C$. Letting $s \to j\omega$, we can express $H(s)$ in the normalized form

$$H(j\omega) = H_0 \frac{j\omega/\omega_0}{1 + j\omega/\omega_0} \tag{3.27a}$$

$$H_0 = -\frac{R_2}{R_1} \qquad \omega_0 = \frac{1}{R_1C} \tag{3.27b}$$

where H_0 is called the *high-frequency gain* and ω_0 is again the -3-dB frequency. As shown in Fig. 3.10*b*, which you are encouraged to justify asymptotically, the circuit is a high-pass filter.

Wideband Band-Pass Filter

The last two circuits can be merged as in Fig. 3.11*a* to give a *band-pass* response. Letting $Z_1 = (R_1C_1s + 1)/C_1s$ and $Z_2 = R_2/(R_2C_2s + 1)$, we get $H(s) = -Z_2/Z_1$, or

$$H(s) = -\frac{R_2}{R_1}\frac{R_1C_1s}{R_1C_1s + 1}\frac{1}{R_2C_2s + 1} \tag{3.28}$$

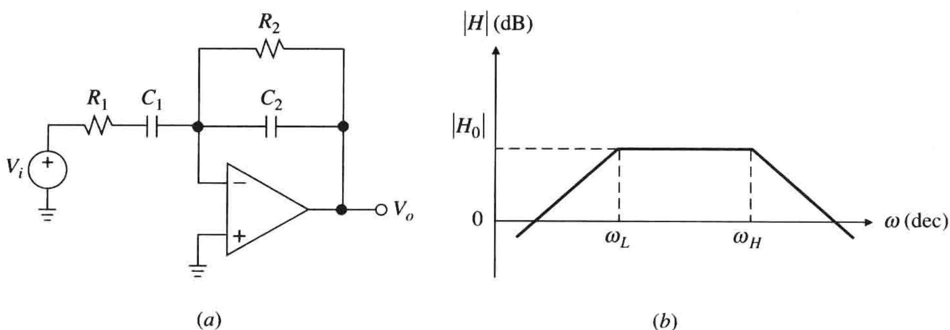


FIGURE 3.11
Wideband band-pass filter.

indicating a zero at the origin and two real poles at $-1/R_1C_1$ and $-1/R_2C_2$. Though this is a second-order filter, we have chosen to discuss it here to demonstrate the use of lower-order building blocks to synthetize higher-order filters. Letting $s \to j\omega$ yields

$$H(j\omega) = H_0 \frac{j\omega/\omega_L}{(1 + j\omega/\omega_L)(1 + j\omega/\omega_H)} \tag{3.29a}$$

$$H_0 = -\frac{R_2}{R_1} \qquad \omega_L = \frac{1}{R_1C_1} \qquad \omega_H = \frac{1}{R_2C_2} \tag{3.29b}$$

where H_0 is called the *midfrequency gain.* The filter is useful with $\omega_L \ll \omega_H$, in which case ω_L and ω_H are called the *low* and *high* -3-dB frequencies. This circuit is used especially in audio applications, where it is desired to amplify signals within the audio range while blocking out subaudio components, such as dc, as well as noise above the audio range.

EXAMPLE 3.5. In the circuit of Fig. 3.11*a* specify suitable component values for a band-pass response with a gain of 20 dB over the audio range.

Solution. For a gain of 20 dB we need $R_2/R_1 = 10$. Try $R_1 = 10$ kΩ and $R_2 = 100$ kΩ. Then, for $\omega_L = 2\pi \times 20$ rad/s we need $C_1 = 1/(2\pi \times 20 \times 10 \times 10^3) = 0.7958$ μF. Use 1 μF, and rescale the resistances as $R_1 = 10^4 \times 0.7958 \cong 7.87$ kΩ and $R_2 = 78.7$ kΩ. For $\omega_H = 2\pi \times 20$ krad/s, use $C_2 = 1/(2\pi \times 20 \times 10^3 \times 78.7 \times 10^3) \cong 100$ pF.

Phase Shifters

In Fig. 3.12*a* the noninverting-input voltage V_p is related to V_i by the low-pass function as $V_p = V_i/(RCs+1)$. Moreover, $V_o = -(R_2/R_1)V_i + (1+R_2/R_1)V_p = 2V_p - V_i$. Eliminating V_p yields

$$H(s) = \frac{-RCs + 1}{RCs + 1} \tag{3.30}$$

indicating a zero at $s = 1/RC$ and a pole at $s = -1/RC$. Letting $s \to j\omega$ yields

$$H(j\omega) = \frac{1 - j\omega/\omega_0}{1 + j\omega/\omega_0} = 1\underline{/-2\tan^{-1}(\omega/\omega_0)} \tag{3.31}$$

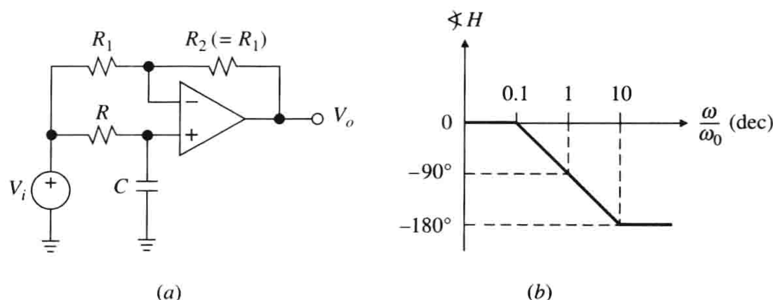


FIGURE 3.12
Phase shifter.

With a gain of 1 V/V, this circuit passes all signals without altering their amplitude. However, as shown in Fig. 3.12*b*, it introduces a variable phase lag from from $0°$ to $-180°$, with a value of $-90°$ at $\omega = \omega_0$. Can you justify using physical insight?

Effect of Finite Open-Loop Gain

As mentioned in the chapter highlights, in this chapter we deliberately assume ideal op amps so we can focus exclusively on filter functions. Though the impact of op amp nonidealities will be investigated in Chapters 5 and 6, we wish to take a quick look at the effect of a noninfinite open-loop gain a to avoid leaving the false impression that op amps do not count. Specifically, we examine the two most basic filters, the integrator and the differentiator (to simplify things further, we assume $r_d = \infty$ and $r_o = 0$).

Viewing the integrator of Fig. 3.6*a* as a V-I converter with C as its load, we adapt Eq. (2.7) to say that C sees an equivalent resistance $R_{eq} = (1 + a)R$. Consequently, the integrator acts as a low-pass filter of the type of Fig. 3.9*a*, with $R_2 = (1 + a)R_1 (\cong aR_1$ for large a). Its unity-gain frequency is still established by R_1 and C; however, Eq. (3.25) indicates also a pole frequency due to aR_1 and C, as well as a dc gain of $-aR_1/R_1 = -a$. (Physically, this makes sense because at sufficiently low frequencies C acts as an open circuit, letting the op amp operate in the open-loop mode).

Turning next to the differentiator of Fig. 3.5*a*, we adapt Eqs. (2.2) and (2.3) to say that C sees an I-V converter with an equivalent input resistance $R_{eq} = R/(1 + a)$. Consequently, the differentiator acts as a high-pass filter of the type of Fig. 3.10*a*, with $R_1 = R_2/(1 + a) (\cong R_2/a$ for large a). Its unity-gain frequency is still established by R_2 and C; however, Eq. (3.27) indicates also a pole frequency due to R_2/a and C, as well as a high-frequency gain of $-R_2/(R_2/a) = -a$.

It is apparent that a noninfinite gain does have an impact on filters. This will be more evident in Chapter 6, where we'll see that the gain a itself is frequency dependent and, as such, it increases the complexity of a filter. But until then, we shall continue assuming ideal op amps for simplicity.

3.3 AUDIO FILTER APPLICATIONS

Audio signal processing provides a multitude of uses for active filters. Common functions required in high-quality audio systems are equalized preamplifiers, active tone control, and graphic equalizers.[6] Equalized preamplifiers are used to compensate for the varying levels at which different parts of the audio spectrum are recorded commercially. Tone control and graphic equalization refer to response adjustments that the listener can effect to compensate for nonideal loudspeaker response, to match apparent room acoustics, or simply to suit one's taste.

Phono Preamplifier

The function of a phono preamplifier is to provide amplification as well as amplitude equalization for the signal from a moving-magnet or a moving-coil cartridge. The response must conform to the standard RIAA (Record Industry Association of America) curve of Fig. 3.13*a*.

Preamplifier gains are usually specified at 1 kHz. The required amount of gain is typically 30 to 40 dB for moving-magnet cartridges, and 50 to 60 dB for moving-coil types. Since the RIAA curve is normalized for unity gain, the actual preamp response will be shifted upward by an amount equal to its gain.

Figure 3.13*b* shows one[7] of several topologies commonly used to approximate the RIAA response. The input-shunting network provides impedance matching for the source, while C_1 provides a low-frequency breakpoint (usually below 20 Hz) to block out dc and any subaudio frequency components. Since $|Z_{C_1}| \ll R_1$ over the frequency range of interest, the transfer function can be found as $H \cong 1 + Z_f/R_1$, where Z_f is the impedance of the feedback network. The result is (see Problem 3.20)

$$H(jf) \cong 1 + \frac{R_2 + R_3}{R_1} \frac{1 + jf/f_1}{(1 + jf/f_2)(1 + jf/f_3)} \tag{3.32}$$

$$f_1 = \frac{1}{2\pi(R_2 \parallel R_3)(C_2 + C_3)} \qquad f_2 = \frac{1}{2\pi R_2 C_2} \qquad f_3 = \frac{1}{2\pi R_3 C_3} \tag{3.33}$$

As long as the circuit is configured for substantially high gain, the unity term in Eq. (3.32) can be ignored, indicating that $H(jf)$ approximates the standard RIAA curve over the audio range.

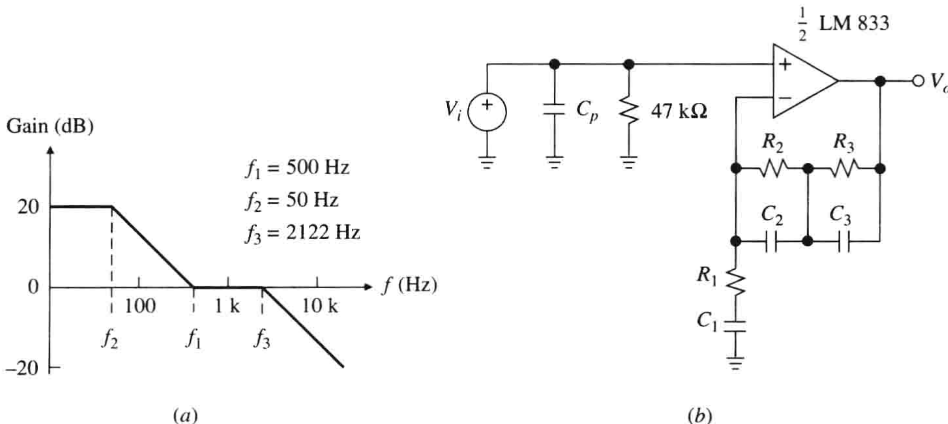


FIGURE 3.13
RIAA playback equalization curve and phono preamplifier.

EXAMPLE 3.6. Design a 40-dB gain, RIAA phono amplifier.

Solution. The RIAA curve must be shifted upward by 40 dB, so the gain below f_2 must be $40 + 20 = 60$ dB $= 10^3$ V/V. Thus, $(R_2 + R_3)/R_1 \cong 10^3$. The expressions for f_1 through f_3 provide three equations in four unknowns. Fix one, say, let $C_2 = 10$ nF. Then, Eq. (3.33) gives $R_2 = 1/(2\pi \times 50 \times 10 \times 10^{-9}) = 318$ kΩ (use 316 kΩ). We also have $1/R_2 + 1/R_3 = 2\pi f_1(C_2 + C_3)$ and $1/R_3 = 2\pi f_3 C_3$. Eliminating $1/R_3$ gives $C_3 = 2.77$ nF (use 2.7 nF). Back substituting gives $R_3 = 27.7$ kΩ (use 28.0 kΩ). Finally, $R_1 = (316 + 28)/10^3 = 344$ Ω (use 340 Ω) and $C_1 = 1/(2\pi \times 340 \times 20) = 23$ μF (use 33 μF). Summarizing, $R_1 = 340$ Ω, $R_2 = 316$ kΩ, $R_3 = 28.0$ kΩ, $C_1 = 33$ μF, $C_2 = 10$ nF, and $C_3 = 2.7$ nF.

Tape Preamplifier

A tape preamplifier must provide gain as well as amplitude and phase equalizations for the signal from a tape head. The response is governed by the standard NAB (National Association of Broadcasters) curve of Fig. 3.14*a*. A circuit[7] to approximate this response is shown in Fig. 3.14*b*. As long as $|Z_{C_1}| \ll R_1$, we have (see Problem 3.18)

$$H(jf) \cong 1 + \frac{R_3}{R_1}\frac{1 + jf/f_1}{1 + jf/f_2} \tag{3.34}$$

$$f_1 = \frac{1}{2\pi R_2 C_2} \qquad f_2 = \frac{1}{2\pi (R_2 + R_3) C_2} \tag{3.35}$$

Active Tone Control

The most common form of tone control is *bass* and *treble* control, which allow the independent adjustment of gain over the lower (bass) and higher (treble) portions

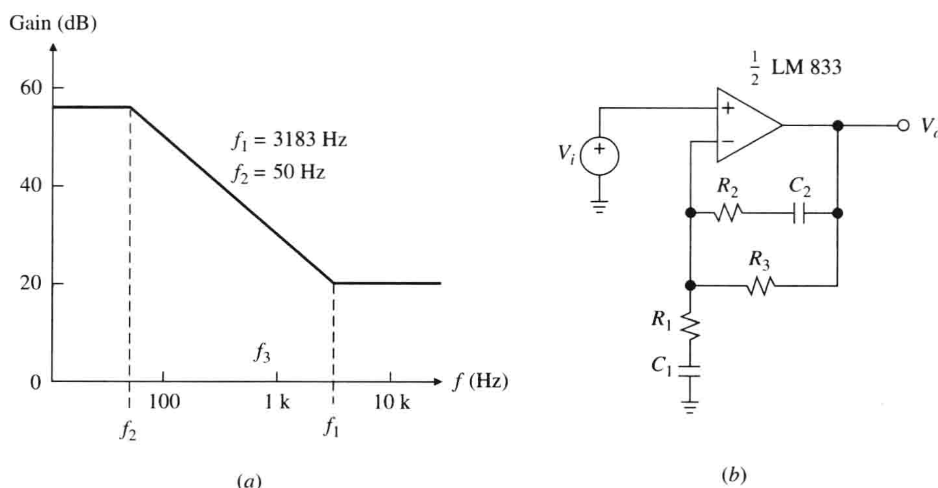
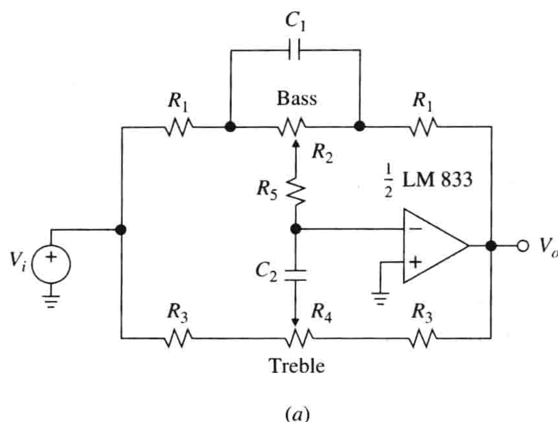
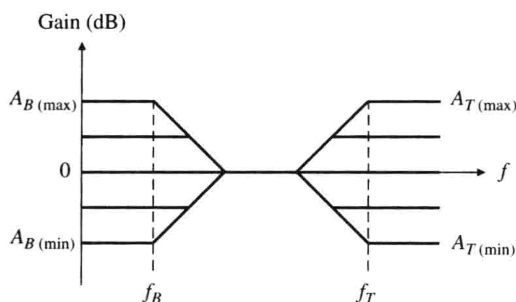


FIGURE 3.14
NAB equalization curve and tape preamplifier.



(a)



(b)

FIGURE 3.15
Bass and treble control.

of the audio range. Figure 3.15 shows one of several circuits in common use and illustrates the effect of tone control on the frequency response.

At the low end of the audio range, or $f < f_B$, the capacitors act as open circuits, so the only effective feedback consists of R_1 and R_2. The op amp acts as an inverting amplifier whose gain magnitude A_B is variable over the range

$$\frac{R_1}{R_1 + R_2} \leq A_B \leq \frac{R_1 + R_2}{R_1} \tag{3.36a}$$

by means of the bass pot. The upper limit is referred to as maximum *boost,* the lower limit as maximum *cut.* For instance, with $R_1 = 11\ \text{k}\Omega$ and $R_2 = 100\ \text{k}\Omega$, these limits are ± 20 dB. Setting the wiper in the middle gives $A_B = 0$ dB, or a *flat* bass response.

As frequency is increased, C_1 gradually bypasses the effect of R_2 until the latter is eventually shorted out and has no effect on the response. The frequency f_B at which C_1 begins to be effective in the case of maximum bass boost or cut is approximately

$$f_B = \frac{1}{2\pi R_2 C_1} \tag{3.36b}$$

Above this frequency the response approaches the flat curve with a slope of about ±6 dB/oct, depending on whether the pot is set for maximum cut or boost.

At the high end of the audio range, or $f > f_T$, the capacitors act as short circuits, so the gain is now controlled by the treble pot. (The bass pot is ineffectual since it is being shorted out by C_1.) It can be proven that if the condition $R_4 \gg (R_1+R_3+2R_5)$ is met, the range of variability of the treble gain A_T is

$$\frac{R_3}{R_1 + R_3 + 2R_5} \leq A_T \leq \frac{R_1 + R_3 + 2R_5}{R_3} \tag{3.37a}$$

and the frequency f_T below which the treble control gradually ceases to affect the response is approximately

$$f_T = \frac{1}{2\pi R_3 C_2} \tag{3.37b}$$

EXAMPLE 3.7. Design a bass/treble control with $f_B = 30$ Hz, $f_T = 10$ kHz, and ±20 dB maximum boost/cut at both ends.

Solution. Since 20 dB corresponds to 10 V/V, we must have $(R_1 + R_2)/R_1 = 10$ and $(R_1 + R_3 + 2R_5)/R_3 = 10$. Let R_2 be a 100-kΩ pot so that $R_1 = 11$ kΩ. Arbitrarily impose $R_5 = R_1 = 11$ kΩ. Then $R_3 = 3.67$ kΩ (use 3.6 kΩ). To meet the condition $R_4 \gg (R_1+R_3+2R_5) \cong 37$ kΩ, let R_4 be a 500-kΩ pot. Then $C_1 = 1/2\pi R_2 f_B = 53$ nF (use 51 nF), and $C_2 = 1/2\pi R_3 f_T = 4.4$ nF (use 5.1 nF). Summarizing, $R_1 = 11$ kΩ, $R_2 = 100$ kΩ, $R_3 = 3.6$ kΩ, $R_4 = 500$ kΩ, $R_5 = 11$ kΩ, $C_1 = 51$ nF, and $C_2 = 5.1$ nF.

Graphic Equalizers

The function of a graphic equalizer is to provide boost and cut control not just at the bass and treble extremes, but also within intermediate frequency bands. Equalizers are implemented with arrays of narrow-band filters whose individual responses are adjusted by vertical slide pots arranged side by side to provide a graphic visualization of the equalized response (hence the name).

Figure 3.16 shows a familiar realization of one of the equalizer sections. The circuit is designed so that over a specified frequency band, C_1 acts as an open circuit while C_2 acts as a short, thus allowing for boost or cut control, depending on whether the wiper position is to the left or to the right, respectively. Outside the band the circuit

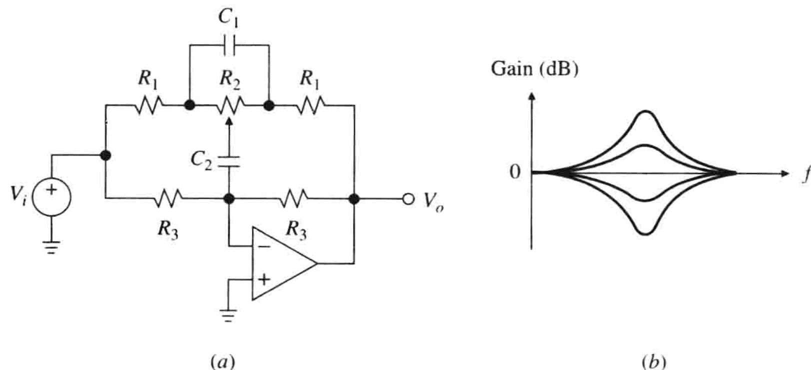


FIGURE 3.16
Section of a graphic equalizer.

provides unity gain, regardless of the wiper position. This stems from the fact that C_2 acts as an open circuit at low frequencies, and C_1 acts as a short at high frequencies. The result is a flat response, but with a peak or a dip over the specified band.

It can be proven[8] that if the component values are chosen so that

$$R_3 \gg R_1 \qquad R_3 = 10R_2 \qquad C_1 = 10C_2 \tag{3.38}$$

then the center of the band is

$$f_0 = \frac{\sqrt{2 + R_2/R_1}}{20\pi R_2 C_2} \tag{3.39a}$$

and the gain magnitude A_0 at this frequency is variable over the range

$$\frac{3R_1}{3R_1 + R_2} \le A_0 \le \frac{3R_1 + R_2}{3R_1} \tag{3.39b}$$

An n-band equalizer is implemented by paralleling n sections and summing the individual outputs with the input in a $1{:}(n-1)$ ratio.[8] This is done with an ordinary summing amplifier. Common choices for the resistances of each section are $R_1 = 10\ \text{k}\Omega$, $R_2 = 100\ \text{k}\Omega$, and $R_3 = 1\ \text{M}\Omega$. The capacitances are calculated using Eqs. (3.38) and (3.39*a*). An equalizer having one section for each octave of the audio spectrum is aptly called an *octave equalizer.*

3.4 STANDARD SECOND-ORDER RESPONSES

Second-order filters are important in their own right and are building blocks of higher-order filters as well, so we investigate their responses in detail before turning to actual circuits.

Looking back at the low-pass, high-pass, and all-pass responses of Section 3.2, we observe that they have the same denominator $D(j\omega) = 1 + j\omega/\omega_0$ and that it is the numerator $N(j\omega)$ that determines the type of response. With $N(j\omega) = 1$ we get the low-pass, with $N(j\omega) = j\omega/\omega_0$ the high-pass, and with $N(j\omega) = 1 - j\omega/\omega_0 = D(j\omega)$ the all-pass response. Moreover, the presence of a scaling factor H_0 does not change the response type; it only shifts its magnitude plot up or down, depending on whether $|H_0| > 1$ or $|H_0| < 1$.

Similar considerations hold for second-order responses. However, since the degree of the denominator is now 2, we have an additional filter parameter besides ω_0. All second-order functions can be put in the standard form

$$H(s) = \frac{N(s)}{(s/\omega_0)^2 + 2\zeta(s/\omega_0) + 1} \tag{3.40}$$

where $N(s)$ is a polynomial in s of degree $m \le 2$; ω_0 is called the *undamped natural frequency,* in radians per second; and ζ (zeta) is a dimensionless parameter called the *damping ratio.* This function has two poles, $p_{1,2} = (-\zeta \pm \sqrt{\zeta^2 - 1})\omega_0$, whose location in the s plane is controlled by ζ as follows:

1. For $\zeta > 1$, the poles are real and negative. The natural response consists of two decaying exponentials and is said to be *overdamped*.
2. For $0 < \zeta < 1$, the poles are complex conjugate and can be expressed as

$$p_{1,2} = -\zeta\omega_0 \pm j\omega_0\sqrt{1 - \zeta^2} \tag{3.41}$$

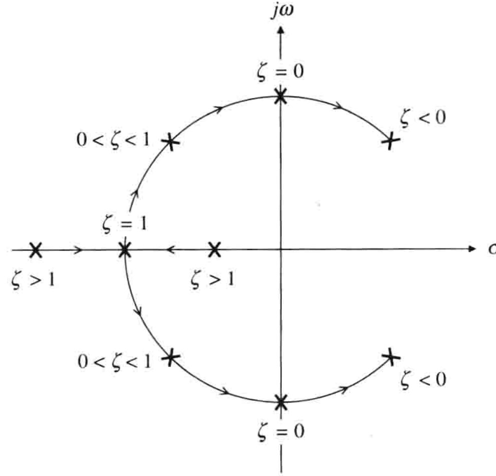


FIGURE 3.17
Root locus for a second-order transfer function.

These poles lie in the left half plane, and the natural response, now called *underdamped,* is the damped sinusoid $x_o(t) = 2|A|e^{-\zeta\omega_0 t}\cos(\omega_0\sqrt{1-\zeta^2}\,t + \measuredangle A)$, where A is the residue at the upper pole.

3. For $\zeta = 0$, Eq. (3.41) yields $p_{1,2} = \pm j\omega_0$, indicating that the poles lie right on the imaginary axis. The natural response is a sustained, or *undamped,* sinusoid with frequency ω_0; hence the name for ω_0.
4. For $\zeta < 0$, the poles lie in the right half plane, thus causing a *diverging* response because the exponent in the term $e^{-\zeta\omega_0 t}$ is now positive. Filters must have $\zeta > 0$ in order to be stable.

The system of trajectories described by the roots as a function of ζ is the root locus depicted in Fig. 3.17. Note that for $\zeta = 1$ the poles are real and coincident.

Letting $s \rightarrow j\omega$ yields the frequency response, which we shall express in terms of the alternative dimensionless parameter Q as

$$H(j\omega) = \frac{N(j\omega)}{1-(\omega/\omega_0)^2 + (j\omega/\omega_0)/Q} \tag{3.42}$$

$$Q = \frac{1}{2\zeta} \tag{3.43}$$

The meaning of Q will become clear as we proceed.

The Low-Pass Response H_{LP}

All second-order low-pass functions can be put in the standard form $H(j\omega) = H_{0LP}H_{LP}(j\omega)$, where H_{0LP} is a suitable constant referred to as the *dc gain,* and

$$H_{LP}(j\omega) = \frac{1}{1-(\omega/\omega_0)^2 + (j\omega/\omega_0)/Q} \tag{3.44}$$

A convenient means for visualizing the frequency plot of H_{LP} (or, for that matter, of any transfer function) is offered by PSpice's Laplace block of Fig. 3.18. The result

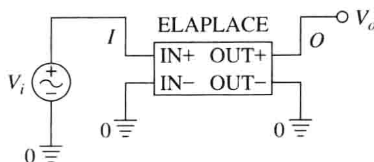


FIGURE 3.18
Using the PSpice's Laplace block to plot $|H_{LP}(j\omega/\omega_0)|_{dB}$ for different values of Q.

is the magnitude plot of Fig. 3.19*a*, with respect to which we make the following considerations:

1. For $\omega/\omega_0 \ll 1$, the second and third denominator terms can be ignored in comparison with unity, so $H_{LP} \to 1$. The low-frequency asymptote is thus

$$|H_{LP}|_{dB} = 0 \qquad (\omega/\omega_0 \ll 1) \tag{3.45a}$$

2. For $\omega/\omega_0 \gg 1$, the second denominator term dominates over the other two, so $H_{LP} \to -1/(\omega/\omega_0)^2$. The high-frequency asymptote is $|H_{LP}|_{dB} = 20\log_{10}[1/(\omega/\omega_0)^2]$, or

$$|H_{LP}|_{dB} = -40\log_{10}(\omega/\omega_0) \qquad (\omega/\omega_0 \gg 1) \tag{3.45b}$$

   This equation is of the type $y = -40x$, or a straight line with a slope of -40 dB/dec. Compared to the first-order response, which has a slope of only -20 dB/dec, the second-order response is closer to the idealized brick-wall profile.
3. For $\omega/\omega_0 = 1$, the two asymptotes meet since letting $\omega/\omega_0 = 1$ in Eq. (3.45*b*) gives Eq. (3.45*a*). Moreover, the first and second denominator terms cancel each other out to give $H_{LP} = -jQ$, or

$$|H_{LP}|_{dB} = Q_{dB} \qquad (\omega/\omega_0 = 1) \tag{3.45c}$$

   In the frequency region near $\omega/\omega_0 = 1$ we now have a family of curves, depending on the value of Q. Contrast this with the first-order case, where only one curve was possible.

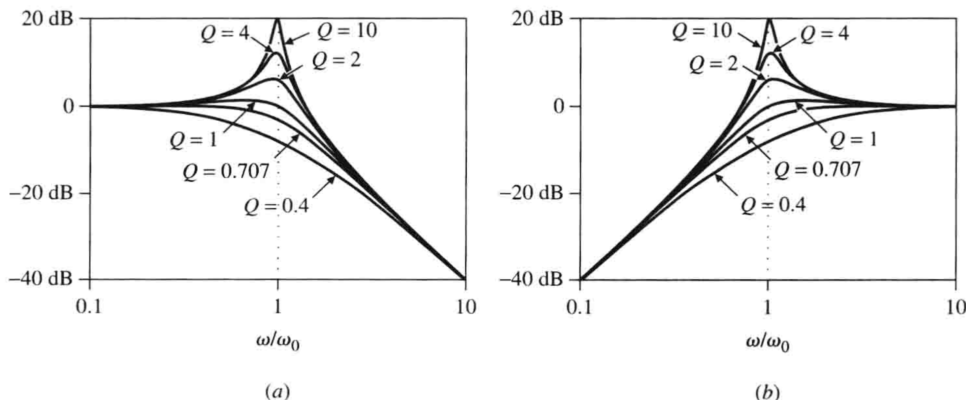


FIGURE 3.19
Magnitude plots for the standard functions (*a*) H_{LP} and (*b*) H_{HP} for different values of Q.

The second-order response, besides providing a high-frequency asymptotic slope twice as steep, offers an additional degree of freedom in specifying the magnitude profile in the vicinity of $\omega/\omega_0 = 1$. In actual applications, Q may range from as low as 0.5 to as high as 100, with values near unity being by far the most common. Note that for low Qs the transition from one asymptote to the other is very gradual, while for high Qs there is a range of frequencies in the vicinity of $\omega/\omega_0 = 1$ where $|H_{LP}| > 1$, a phenomenon referred to as *peaking.*

One can prove that the largest Q before the onset of peaking is $Q = 1/\sqrt{2} = 0.707$. The corresponding curve is said to be *maximally flat* and is also referred to as the *Butterworth response.* This curve is the closest to the brick-wall model, hence its widespread use. By Eq. (3.45*c*), $|H_{LP}|_{dB} = (1/\sqrt{2})_{dB} = -3$ dB. The meaning of ω_0 for the Butterworth response is the same as for the first-order case, that is, ω_0 represents the *−3-dB frequency,* also called the *cutoff frequency.*

It can be proven[5] that in the case of peaked responses, or $Q > 1/\sqrt{2}$, the frequency at which $|H_{LP}|$ is maximized and the corresponding maximum are

$$\omega/\omega_0 = \sqrt{1 - 1/2Q^2} \tag{3.46a}$$

$$|H_{LP}|_{max} = \frac{Q}{\sqrt{1 - 1/4Q^2}} \tag{3.46b}$$

For sufficiently large Qs, say, $Q > 5$, we have $\omega/\omega_0 \cong 1$ and $|H_{LP}|_{max} \cong Q$. Of course, in the absence of peaking, or $Q < 1/\sqrt{2}$, the maximum is reached at $\omega/\omega_0 = 0$, that is, at dc. Peaked responses are useful in the cascade synthesis of higher-order filters, to be covered in Chapter 4.

The High-Pass Response H_{HP}

The standard form of all second-order high-pass functions is $H(j\omega) = H_{0HP}H_{HP}(j\omega)$, where H_{0HP} is called the *high-frequency gain,* and

$$H_{HP}(j\omega) = \frac{-(\omega/\omega_0)^2}{1 - (\omega/\omega_0)^2 + (j\omega/\omega_0)/Q} \tag{3.47}$$

(Note that the negative sign in the numerator is part of the definition.) Letting $j\omega \to s$ reveals that $H(s)$, besides the pole pair, has a double zero at the origin. To plot $|H_{HP}(j\omega/\omega_0)|$ we use again the Laplace block of Fig. 3.18, but with the quadratic term in s at the numerator. The result, shown in Fig. 3.19*b*, is the mirror image of that of H_{LP}. This symmetry stems from the (readily verifiable) fact that the function $H_{HP}(j\omega/\omega_0)$ can be obtained from $H_{LP}(j\omega/\omega_0)$ by replacing $(j\omega/\omega_0)$ with $1/(j\omega/\omega_0)$.

The Band-Pass Response H_{BP}

The standard form of all second-order band-pass functions is $H(j\omega) = H_{0BP}H_{BP}(j\omega)$, where H_{0BP} is the called the *resonance gain,* and

$$H_{BP}(j\omega) = \frac{(j\omega/\omega_0)/Q}{1 - (\omega/\omega_0)^2 + (j\omega/\omega_0)/Q} \tag{3.48}$$

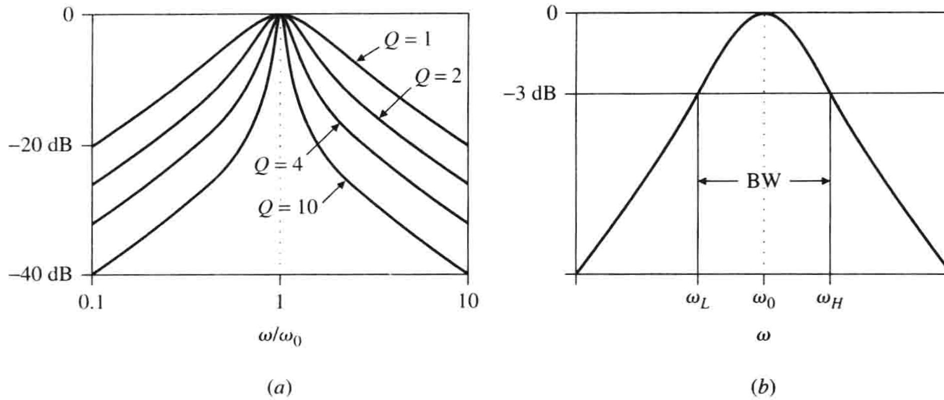


FIGURE 3.20
(*a*) Magnitude plots of the standard function H_{BP} for different values of Q. (*b*) Visualizing the −3-dB bandwidth BW.

(Note that Q in the numerator is part of the definition.) Besides the pole pair, this function has a zero at the origin. To plot $|H_{BP}(j\omega/\omega_0)|$, we use again the Laplace block of Fig. 3.18, but with the term $(s/6.283)/Q$ as the numerator. The result is shown in Fig. 3.20*a*, with respect to which we make the following considerations:

1. For $\omega/\omega_0 \ll 1$, we can ignore the second and third denominator terms and write $H_{BP} \to (j\omega/\omega_0)/Q$. The low-frequency asymptote is thus $|H_{BP}|_{dB} = 20\log_{10}[(\omega/\omega_0)/Q]$, or

$$|H_{BP}|_{dB} = 20\log_{10}(\omega/\omega_0) - Q_{dB} \qquad (\omega/\omega_0 \ll 1) \qquad (3.49a)$$

   This equation is of the type $y = 20x - Q_{dB}$, indicating a straight line with a slope of +20 dB/dec, but shifted by $-Q_{dB}$ with respect to the 0-dB axis at $\omega/\omega_0 = 1$.
2. For $\omega/\omega_0 \gg 1$, the second term dominates in the denominator, so $H_{BP} \to -j1/(\omega/\omega_0)Q$. The high-frequency asymptote is thus

$$|H_{BP}|_{dB} = -20\log_{10}(\omega/\omega_0) - Q_{dB} \qquad (\omega/\omega_0 \gg 1) \qquad (3.49b)$$

   This is a straight line with the same amount of downshift as before, but with a slope of −20 dB/dec.
3. For $\omega/\omega_0 = 1$, we get $H_{BP} = 1$, or

$$|H_{BP}|_{dB} = 0 \qquad (\omega/\omega_0 = 1) \qquad (3.49c)$$

   One can prove that $|H_{BP}|$ peaks at $\omega/\omega_0 = 1$ regardless of Q, this being the reason why ω_0 is called the *peak,* or *resonance, frequency.*

Figure 3.20*a* indicates that all curves peak at 0 dB. Those corresponding to low Qs are broad, but those corresponding to high Qs are narrow, indicating a higher degree of selectivity. In the vicinity of $\omega/\omega_0 = 1$ the high-selectivity curves are much steeper than ±20 dB/dec, though away from resonance they roll off at the same ultimate rate of ±20 dB/dec.

To express selectivity quantitatively, we introduce the *bandwidth*

$$\mathrm{BW} = \omega_H - \omega_L \tag{3.50}$$

where ω_L and ω_H are the -3-dB frequencies, that is, the frequencies at which the response is 3 dB below its maximum, as depicted in Fig. 3.20*b*. One can prove[5] that

$$\omega_L = \omega_0\left(\sqrt{1 + 1/4Q^2} - 1/2Q\right) \tag{3.51a}$$

$$\omega_H = \omega_0\left(\sqrt{1 + 1/4Q^2} + 1/2Q\right) \tag{3.51b}$$

$$\omega_0 = \sqrt{\omega_L \omega_H} \tag{3.52}$$

The resonance frequency ω_0 is the *geometric mean* of ω_L and ω_H, indicating that on a logarithmic scale ω_0 appears halfway between ω_L and ω_H. It is apparent that the narrower the bandwidth, the more selective the filter. However, selectivity depends also on ω_0, since a filter with BW = 10 rad/s and $\omega_0 = 1$ krad/s is certainly more selective than one with BW = 10 rad/s but $\omega_0 = 100$ rad/s. A proper measure of selectivity is the ratio ω_0/BW. Subtracting Eq. (3.51*a*) from Eq. (3.51*b*) and taking the reciprocal, we get

$$Q = \frac{\omega_0}{\mathrm{BW}} \tag{3.53}$$

that is, Q *is* the selectivity. We now have a more concrete interpretation for this parameter.

The Notch Response H_N

The most common form for the notch function is $H(j\omega) = H_{0\mathrm{N}} H_\mathrm{N}(j\omega)$, where $H_{0\mathrm{N}}$ is an appropriate gain constant, and

$$H_\mathrm{N}(j\omega) = \frac{1 - (\omega/\omega_0)^2}{1 - (\omega/\omega_0)^2 + (j\omega/\omega_0)/Q} \tag{3.54}$$

(In Section 3.7 we shall see that other notch functions are possible, in which ω_0 in the numerator has not necessarily the same value as ω_0 in the denominator.) Letting $j\omega \to s$ reveals that $H(s)$, besides the pole pair, has a zero pair on the imaginary axis, or $z_{1,2} = \pm j\omega_0$. We observe that at sufficiently low and high frequencies, $H_\mathrm{N} \to 1$. However, for $\omega/\omega_0 = 1$ we get $H_\mathrm{N} \to 0$, or $|H_\mathrm{N}|_{\mathrm{dB}} \to -\infty$. To plot $|H_N(j\omega/\omega_0)|$ we use again the Laplace block of Fig. 3.18, but with $1+(s/6.283)**2$ as the numerator. The result, shown in Fig. 3.21*a*, indicates that the higher the Q the narrower the notch. For obvious reasons, ω_0 is called the *notch frequency.* In a practical circuit, because of component nonidealities, an infinitely deep notch is unrealizable.

It is interesting to note that

$$H_\mathrm{N} = H_\mathrm{LP} + H_\mathrm{HP} = 1 - H_\mathrm{BP} \tag{3.55}$$

indicating alternative ways of synthesizing the notch response once the other responses are available.

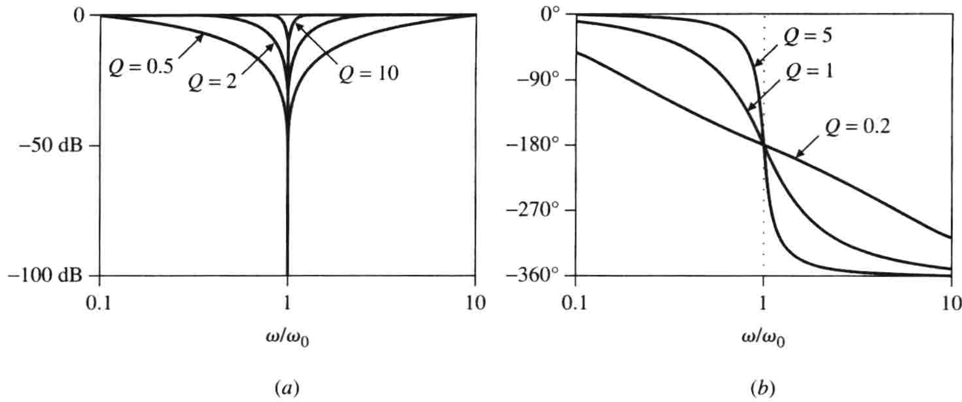


FIGURE 3.21
(a) Magnitude plot of H_N and (b) phase plot of H_{AP}, for different values of Q.

The All-Pass Response H_{AP}

Its general form is $H(j\omega) = H_{0AP}H_{AP}(j\omega)$, where H_{0AP} is the usual gain term, and

$$H_{AP}(j\omega) = \frac{1 - (\omega/\omega_0)^2 - (j\omega/\omega_0)/Q}{1 - (\omega/\omega_0)^2 + (j\omega/\omega_0)/Q} \tag{3.56}$$

This function has two poles and two zeros. For $Q > 0.5$, the zeros and poles are complex and are symmetrical about the $j\omega$ axis. Since $N(j\omega) = D(j\omega)$, we have $|H_{AP}| = 1$, or $|H_{AP}|_{dB} = 0$ dB, regardless of frequency. The argument is

$$\sphericalangle H_{AP} = -2\tan^{-1}\frac{(\omega/\omega_0)/Q}{1 - (\omega/\omega_0)^2} \qquad \text{for } \omega/\omega_0 < 1 \tag{3.57a}$$

$$\sphericalangle H_{AP} = -360^\circ - 2\tan^{-1}\frac{(\omega/\omega_0)/Q}{1 - (\omega/\omega_0)^2} \qquad \text{for } \omega/\omega_0 > 1 \tag{3.57b}$$

indicating that as ω/ω_0 is swept from 0 to ∞, the argument changes from 0°, through −180°, to −360°. This is shown in Fig. 3.21b. The all-pass function can also be synthesized as

$$H_{AP} = H_{LP} - H_{BP} + H_{HP} = 1 - 2H_{BP} \tag{3.58}$$

Filter Measurements

Because of component tolerances and other nonidealities, the parameters of a practical filter are likely to deviate from their design values. We thus need to measure them and, if necessary, to tune them via suitable potentiometers.

For a low-pass filter we have $H_{LP}(j0) = H_{0LP}$ and $H_{LP}(j\omega_0) = -jH_{0LP}Q$. To measure ω_0, we look for the frequency at which the output is shifted by 90° with respect to the input, and to measure Q, we take the ratio $Q = |H_{LP}(j\omega_0)|/|H_{0LP}|$.

For a band-pass filter we have $H_{BP}(j\omega_0) = H_{0BP}$, $\measuredangle H_{BP}(j\omega_L) = \measuredangle H_{0BP} - 45°$, and $\measuredangle H_{BP}(j\omega_H) = \measuredangle H_{0BP} - 135°$. Thus, ω_0 is measured as the frequency at which the output is in phase with the input if $H_{0BP} > 0$, or 180° out of phase if $H_{0BP} < 0$. To find Q, we measure the frequencies ω_L and ω_H at which the output is shifted by ±45° with respect to the input. Then, $Q = \omega_0/(\omega_H - \omega_L)$. The reader can apply similar considerations to measure the parameters of the other responses.

3.5 *KRC* FILTERS

Since an *R-C* stage provides a first-order low-pass response, cascading two such stages as in Fig. 3.22*a* ought to provide a second-order response, and without using any inductances. Indeed, at low frequencies the capacitors act as open circuits, thus letting the input signal pass through with $H \to 1$ V/V. At high frequencies the incoming signal will be shunted to ground first by C_1 and then by C_2, thus providing a two-step attenuation; hence the designation *second-order.* Since at high frequencies a single *R-C* stage gives $H \to 1/(j\omega/\omega_0)$, the cascade combination of two stages gives $H \to [1/(j\omega/\omega_1)] \times [1/(j\omega/\omega_2)] = -1/(\omega/\omega_0)^2$, $\omega_0 = \sqrt{\omega_1\omega_2}$, indicating an asymptotic slope of −40 dB/dec. The filter of Fig. 3.22*a* does meet the asymptotic criteria for a second-order low-pass response; however, it does not offer sufficient flexibility for controlling the magnitude profile in the vicinity of $\omega/\omega_0 = 1$. In fact, one can prove[5] that this all-passive filter yields $Q < 0.5$.

If we wish to increase Q above 0.5, we must bolster the magnitude response near $\omega = \omega_0$. One way to achieve this is by providing a controlled amount of *positive feedback.* In Fig. 3.22*b* the output of the R_2–C_2 stage is magnified by an amplifier with gain K, and then is fed back to the interstage node via C_1, whose bottom terminal has been lifted off ground to create the positive feedback path. This feedback must be effective only in the vicinity of $\omega = \omega_0$, where bolstering is specifically needed. We can use physical insight to verify the band-pass nature of the feedback: for $\omega/\omega_0 \ll 1$ the impedance of C_1 is simply too large to feed back much signal, whereas for $\omega/\omega_0 \gg 1$, the shunting action by C_2 makes V_o too small to do much good. However, near $\omega/\omega_0 = 1$ there will be feedback, which we can adjust for the desired amount of peaking by acting on K. Filters of the type of Fig. 3.22*b* are aptly called *KRC* filters—or also *Sallen-Key filters,* for their inventors.

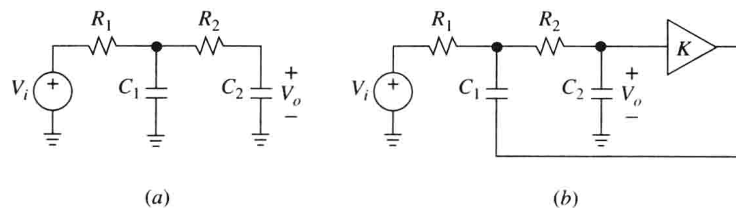


FIGURE 3.22
(*a*) Passive and (*b*) active realization of a second-order low-pass filter.

Low-Pass *KRC* Filters

In Fig. 3.23 the gain block is implemented with an op amp operating as a noninverting amplifier, and

$$K = 1 + \frac{R_B}{R_A} \tag{3.59}$$

Note that V_o is obtained from the output node of the op amp to take advantage of its low impedance. By inspection,

$$V_o = K \frac{1}{R_2 C_2 s + 1} V_1$$

Summing currents at node V_1,

$$\frac{V_i - V_1}{R_1} + \frac{V_o/K - V_1}{R_2} + \frac{V_o - V_1}{1/C_1 s} = 0$$

Eliminating V_1 and collecting, we get

$$H(s) = \frac{V_o}{V_i} = \frac{K}{R_1 C_1 R_2 C_2 s^2 + [(1-K)R_1 C_1 + R_1 C_2 + R_2 C_2]s + 1}$$

Letting $s \to j\omega$ yields

$$H(j\omega) = K \frac{1}{1 - \omega^2 R_1 C_1 R_2 C_2 + j\omega[(1-K)R_1 C_1 + R_1 C_2 + R_2 C_2]}$$

Next, we put this function in the standard form $H(j\omega) = H_{0LP} H_{LP}(j\omega)$, with $H_{LP}(j\omega)$ as in Eq. (3.44). To do so, we equate the coefficients pairwise. By inspection,

$$H_{0LP} = K \tag{3.60a}$$

Letting $\omega^2 R_1 C_1 R_2 C_2 = (\omega/\omega_0)^2$ gives

$$\omega_0 = \frac{1}{\sqrt{R_1 C_1 R_2 C_2}} \tag{3.60b}$$

indicating that ω_0 is the geometric mean of the individual-stage frequencies $\omega_1 = 1/R_1 C_1$ and $\omega_2 = 1/R_2 C_2$. Finally, letting $j\omega[(1-K)R_1 C_1 + R_1 C_2 + R_2 C_2] =$

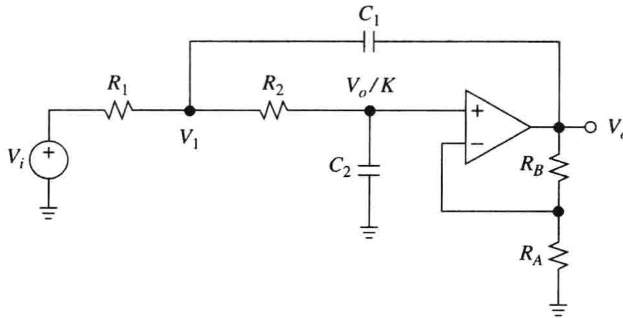


FIGURE 3.23
Low-pass *KRC* filter.

$(j\omega/\omega_0)/Q$ gives

$$Q = \frac{1}{(1-K)\sqrt{R_1C_1/R_2C_2} + \sqrt{R_1C_2/R_2C_1} + \sqrt{R_2C_2/R_1C_1}} \tag{3.60c}$$

We observe that K and Q depend on component *ratios,* while ω_0 depends on component *products.* Because of component tolerances and op amp nonidealities, the parameters of an actual filter are likely to depart from their intended values. Our filter can be tuned as follows: (a) adjust R_1 for the desired ω_0 (this adjustment varies Q also); (b) once ω_0 has been tuned, adjust R_B for the desired Q (this leaves ω_0 unchanged; however, it varies K, but this is of little concern because it does not affect the frequency behavior).

Since we have five parameters (K, R_1, C_1, R_2, and C_2) but only three equations, we have the choice of fixing two so we can specify design equations for the remaining three. Two common designs are the *equal-component* and the *unity-gain* designs (other designs are discussed in the end-of-chapter problems).

Equal-Component *KRC* Circuit

Imposing $R_1 = R_2 = R$ and $C_1 = C_2 = C$ simplifies inventory and reduces Eq. (3.60) to

$$H_{0\mathrm{LP}} = K \qquad \omega_0 = \frac{1}{RC} \qquad Q = \frac{1}{3-K} \tag{3.61}$$

The design equations are then

$$RC = 1/\omega_0 \qquad K = 3 - 1/Q \qquad R_B = (K-1)R_A \tag{3.62}$$

EXAMPLE 3.8. Using the equal-component design, specify elements for a second-order low-pass filter with $f_0 = 1$ kHz and $Q = 5$. What is its dc gain?

Solution. Arbitrarily select $C = 10$ nF, which is an easily available value. Then, $R = 1/(\omega_0 C) = 1/(2\pi 10^3 \times 10 \times 10^{-9}) = 15.92$ kΩ (use 15.8 kΩ, 1%). Moreover, $K = 3 - 1/5 = 2.80$, and $R_B/R_A = 2.80 - 1 = 1.80$. Let $R_A = 10.0$ kΩ, 1%; then, $R_B = 17.8$ kΩ, 1%. The circuit, shown in Fig. 3.24*a*, has a dc gain of 2.78 V/V.

EXAMPLE 3.9. Modify the circuit of Example 3.8 for a dc gain of 0 dB.

Solution. This situation arises often enough to merit a detailed treatment. To reduce gain from an existing value A_{old} to a different value A_{new}, apply Thévenin's theorem and replace R_1 with a voltage divider R_{1A} and R_{1B} such that

$$A_{\mathrm{new}} = \frac{R_{1B}}{R_{1A} + R_{1B}} A_{\mathrm{old}} \qquad R_{1A} \parallel R_{1B} = R_1$$

where the second constraint ensures that ω_0 is unaffected by the replacement. Solving, we get

$$R_{1A} = R_1 \frac{A_{\mathrm{old}}}{A_{\mathrm{new}}} \qquad R_{1B} = \frac{R_1}{1 - A_{\mathrm{new}}/A_{\mathrm{old}}} \tag{3.63}$$

In our case, $A_{\mathrm{old}} = 2.8$ V/V and $A_{\mathrm{new}} = 1$ V/V. So, $R_{1A} = 15.92 \times 2.8/1 = 44.56$ kΩ (use 44.2 kΩ, 1%) and $R_{1B} = 15.92/(1 - 1/2.8) = 24.76$ kΩ (use 24.9 kΩ, 1%). The circuit is shown in Fig. 3.24*b*.

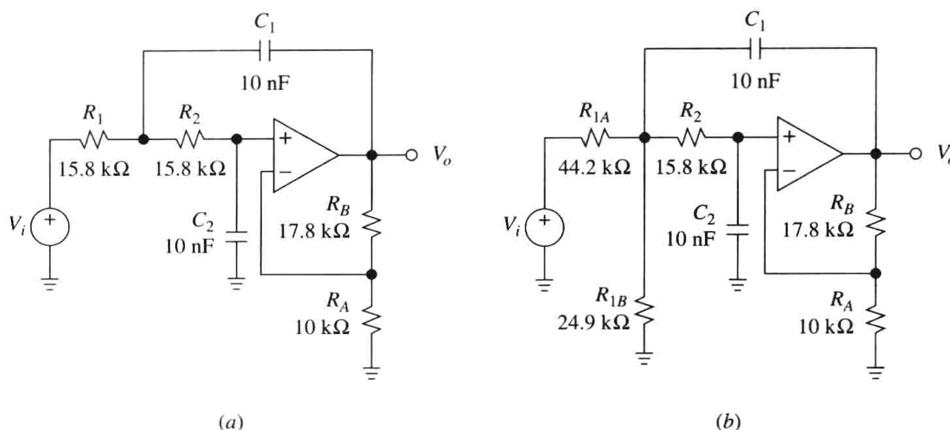


FIGURE 3.24
Filter realizations of Examples 3.8 and 3.9.

Unity-Gain *KRC* Circuit

Imposing $K = 1$ minimizes the number of components and also maximizes the bandwidth of the op amp, an issue that will be studied in Chapter 6. To simplify the math, we relabel the components as $R_2 = R$, $C_2 = C$, $R_1 = mR$, and $C_1 = nC$. Then, Eq. (3.60) reduces to

$$H_{0\text{LP}} = 1\text{ V/V} \qquad \omega_0 = \frac{1}{\sqrt{mn}RC} \qquad Q = \frac{\sqrt{mn}}{m+1} \tag{3.64}$$

You can verify that for a given n, Q is maximized when $m = 1$, that is, when the resistances are equal. With $m = 1$, Eq. (3.64) gives $n = 4Q^2$. In practice, one starts out with two easily available capacitances in a ratio $n \geq 4Q^2$; then m is found as $m = k + \sqrt{k^2 - 1}$, where $k = n/2Q^2 - 1$.

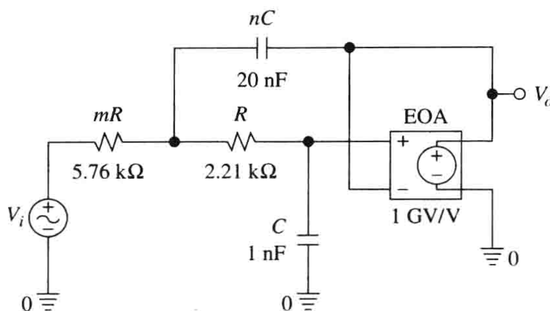
EXAMPLE 3.10. (*a*) Using the unity-gain option, design a low-pass filter with $f_0 =$ 10 kHz and $Q = 2$. (*b*) Use PSpice to visualize its frequency response, both magnitude and phase.

Solution.

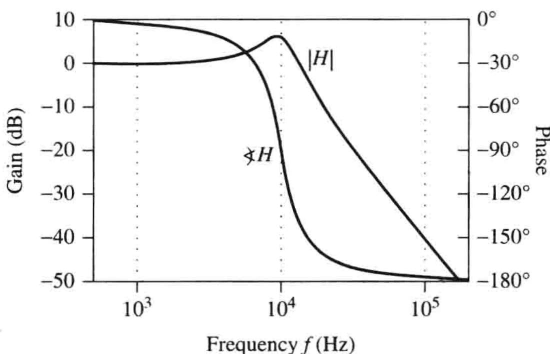
(*a*) Arbitrarily pick $C = 1$ nF. Since $4Q^2 = 4 \times 2^2 = 16$, let $n = 20$. Then, $nC = 20$ nF, $k = 20/(2 \times 2^2) - 1 = 1.5$, $m = 1.5 + \sqrt{1.5^2 - 1} = 2.618$, $R = 1/(\sqrt{mn}\omega_0 C) = 1/(\sqrt{2.618 \times 20} \times 2\pi 10^4 \times 10^{-9}) = 2.199$ kΩ (use 2.21 kΩ, 1%), and $mR =$ 5.758 kΩ (use 5.76 kΩ, 1%). The filter is shown in Fig. 3.25.

(*b*) Using the PSpice circuit of Fig. 3.25*a*, we generate the Bode plots of Fig. 3.25*b*.

EXAMPLE 3.11. (*a*) Design a second-order low-pass Butterworth filter with a −3-dB frequency of 10 kHz. (*b*) If $v_i(t) = 10\cos(4\pi 10^4 t + 90°)$ V, find $v_o(t)$.



(a)



(b)

FIGURE 3.25
(a) PSpice circuit for the low-pass filter of Example 3.10, and (b) its Bode plots.

Solution.

(a) The Butterworth response, for which $Q = 1/\sqrt{2}$, is implemented with $m = 1$ and $n = 2$. Letting $C = 1$ nF, we get $nC = 2$ nF and $mR = R = 11.25$ kΩ (use 11.3 kΩ, 1%).

(b) Since $\omega/\omega_0 = 2$, we have $H(j4\pi 10^4) = 1/[1 - 2^2 + j2/(1/\sqrt{2})] = (1/\sqrt{17})\,\underline{/-136.69^\circ}$ V/V. So, $V_{om} = 10/\sqrt{17} = 2.426$ V, $\theta_o = -136.69^\circ + 90^\circ = -46.69^\circ$, and $v_o(t) = 2.426 \cos(4\pi 10^4 t - 46.69^\circ)$ V.

The advantages of the unity-gain design are offset by a quadratic increase of the capacitance spread n with Q. Moreover, the circuit does not enjoy the tuning advantages of the equal-component design because the adjustments of ω_0 and Q interfere with each other, as revealed by Eq. (3.64). On the other hand, at high Qs the equal-component design becomes too sensitive to the tolerances of R_B and R_A, when their ratio is very close to 2. A slight mismatch may cause an intolerable departure of Q from the desired value. Should this ratio reach (or even surpass) 2, Q will become infinite (or even negative), causing the filter to oscillate. For these reasons, *KRC* filters are used for Qs below 10. Section 3.7 presents filter topologies suited to high Qs.

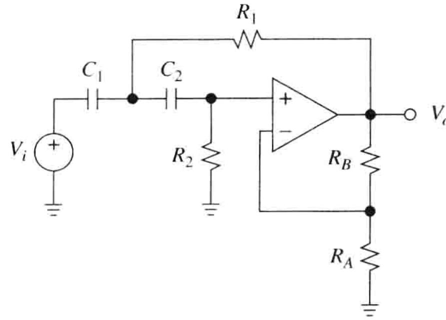


FIGURE 3.26
High-pass *KRC* filter.

High-Pass *KRC* Filters

Interchanging the components of a low-pass *R-C* stage with each other turns it into a high-pass *C-R* stage. Interchanging resistances and capacitances in the low-pass filter of Fig. 3.23 leads to the filter of Fig. 3.26, which you can readily classify as a high-pass type using physical insight. By similar analysis, we find that $V_o/V_i = H_{0\text{HP}} H_{\text{HP}}$, where H_{HP} is given in Eq. (3.47), and

$$H_{0\text{HP}} = K \qquad \omega_0 = \frac{1}{\sqrt{R_1 C_1 R_2 C_2}} \tag{3.65a}$$

$$Q = \frac{1}{(1-K)\sqrt{R_2C_2/R_1C_1} + \sqrt{R_1C_2/R_2C_1} + \sqrt{R_1C_1/R_2C_2}} \tag{3.65b}$$

As in the low-pass case, two interesting options available to the designer are the *equal-component* and the *unity-gain* designs.

EXERCISE 3.1. Derive Eq. (3.65).

EXAMPLE 3.12. Design a second-order high-pass filter with $f_0 = 200$ Hz and $Q = 1.5$.

Solution. To minimize the component count, choose the unity-gain option, for which $R_A = \infty$ and $R_B = 0$. Letting $C_1 = nC_2$ and $R_1 = mR_2$ in Eq. (3.65) gives $\omega_0 = 1/\sqrt{mn}RC$ and $Q = (\sqrt{n/m})/(n+1)$. Let $C_1 = C_2 = 0.1\ \mu$F, so that $n = 1$. Imposing $1.5 = (\sqrt{1/m})/2$ gives $m = 1/9$, and imposing $2\pi 200 = 1/(\sqrt{1/9}R_2 \times 10^7)$ gives $R_2 = 23.87$ kΩ and $R_1 = mR_2 = 2.653$ kΩ.

Band-Pass *KRC* Filters

The circuit of Fig. 3.27 consists of an *R-C* stage followed by a *C-R* stage to synthesize a band-pass block, and a gain block to provide positive feedback via R_3. This feedback is designed to bolster the response near $\omega/\omega_0 = 1$. The ac analysis of the filter yields $V_o/V_i = H_{0\text{BP}} H_{\text{BP}}$, where H_{BP} is given in Eq. (3.48), and

$$H_{0\text{BP}} = \frac{K}{1 + (1-K)R_1/R_3 + (1 + C_1/C_2)R_1/R_2} \qquad \omega_0 = \frac{\sqrt{1 + R_1/R_3}}{\sqrt{R_1C_1R_2C_2}} \tag{3.66a}$$

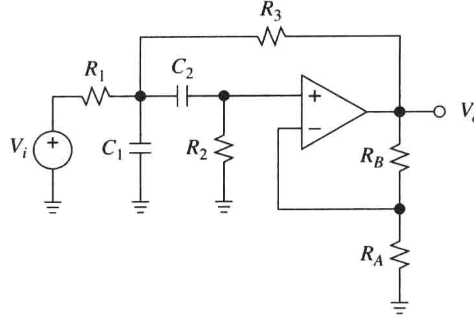


FIGURE 3.27
Band-pass *KRC* filter.

$$Q = \frac{\sqrt{1 + R_1/R_3}}{[1 + (1 - K)R_1/R_3]\sqrt{R_2C_2/R_1C_1} + \sqrt{R_1C_2/R_2C_1} + \sqrt{R_1C_1/R_2C_2}} \tag{3.66b}$$

We again note that one can vary R_1 to tune ω_0 and R_B to adjust Q.

If $Q > \sqrt{2}/3$, a convenient choice is $R_1 = R_2 = R_3 = R$ and $C_1 = C_2 = C$, in which case the above expressions reduce to

$$H_{0\mathrm{BP}} = \frac{K}{4 - K} \qquad \omega_0 = \frac{\sqrt{2}}{RC} \qquad Q = \frac{\sqrt{2}}{4 - K} \tag{3.67}$$

The corresponding design equations are

$$RC = \sqrt{2}/\omega_0 \qquad K = 4 - \sqrt{2}/Q \qquad R_B = (K - 1)R_A \tag{3.68}$$

EXERCISE 3.2. Derive Eqs. (3.66) through (3.68).

EXAMPLE 3.13. (*a*) Design a second-order band-pass filter with $f_0 = 1$ kHz and BW $= 100$ Hz. What is its resonance gain? (*b*) Modify the circuit for a resonance gain of 20 dB.

Solution.

(*a*) Use the equal-component option with $C_1 = C_2 = 10$ nF and $R_1 = R_2 = R_3 = \sqrt{2}/(2\pi 10^3 \times 10^{-8}) = 22.5$ kΩ (use 22.6 kΩ, 1%). We need $Q = f_0/\mathrm{BW} = 10$, so $K = 4 - \sqrt{2}/10 = 3.858$. Pick $R_A = 10.0$ kΩ, 1%. Then, $R_B = (K - 1)R_A = 28.58$ kΩ (use 28.7 kΩ, 1%). The resonance gain is $K/(4 - K) = 27.28$ V/V.

(*b*) Replace R_1 with two resistances R_{1A} and R_{1B}, in the manner of Example 3.9, whose values are found via Eq. (3.63) with $A_{\mathrm{old}} = 27.28$ V/V and $A_{\mathrm{new}} = 10^{20/20} = 10$ V/V. This gives $R_{1A} = 61.9$ kΩ, 1%, and $R_{1B} = 35.7$ kΩ, 1%.

Band-Reject *KRC* Filters

The circuit of Fig. 3.28 consists of a *twin-T*-network and a gain block to provide positive feedback via the top capacitance. The *T*-networks provide alternative forward paths through which V_i can reach the amplifier's input: the low-frequency path *R-R*, and the high-frequency path *C-C*, indicating $H \rightarrow K$ at the frequency extremes. At intermediate frequencies, however, the two paths provide opposing phase angles,

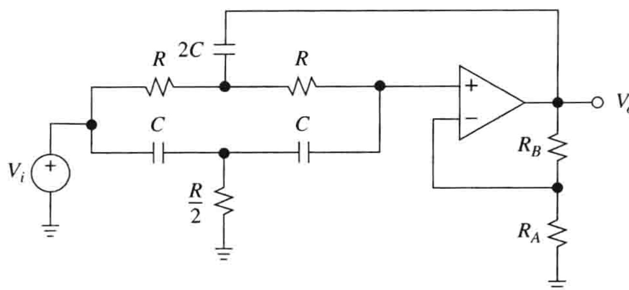


FIGURE 3.28
Band-reject *KRC* filter.

indicating a tendency of the two forward signals to cancel each other out at the amplifier's input. We thus anticipate a notch response. The ac analysis of the circuit gives $V_o/V_i = H_{0N}H_N$, where H_N is given in Eq. (3.54), and

$$H_{0N} = K \qquad \omega_0 = \frac{1}{RC} \qquad Q = \frac{1}{4-2K} \tag{3.69}$$

EXERCISE 3.3. Derive Eq. (3.69).

EXAMPLE 3.14. (*a*) Using standard 1% components, design a notch filter having $f_0 =$ 60 Hz and BW = 5 Hz. (*b*) Use PSpice to measure the actual notch frequency and depth of the notch, and comment.

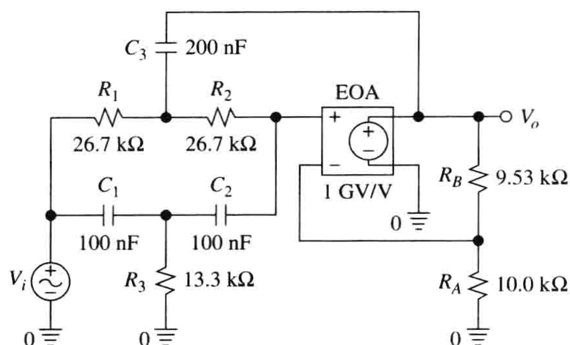
Solution.

(*a*) Let $C = 100$ nF and $2C = 200$ nF. Then, $R = 1/(2\pi 60 \times 10^{-7}) = 26.5258$ kΩ (use 26.7 kΩ, 1%), and $R/2 = 13.2629$ kΩ (use 13.3 kΩ, 1%). Since $Q = 60/5 = 12$, we get $K = 4 - 1/12 = 47/24$, or $R_B/R_A = 23/24$. Use $R_A = 10.0$ kΩ and $R_B = 9.53$ kΩ, both 1%.

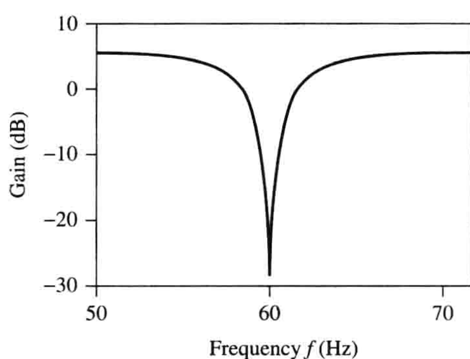
(*b*) Using the circuit of Fig. 3.29*a* we get the magnitude plot of Fig. 3.29*b*. Cursor measurements give a notch frequency of $f_0 = 59.665$ Hz and a depth of $|H_{N(\min)}(jf_0)| = -28.4$ dB. The departure from ideality is due primarily to the usage of 1% component values. Rerunning PSpice with the *calculated* (as opposed to 1%) component values, we get $f_0 = 60.000$ Hz and $|H_{N(\min)}(jf_0)| = -75$ dB.

3.6 MULTIPLE-FEEDBACK FILTERS

Multiple-feedback filters utilize more than one feedback path. Unlike their *KRC* counterparts, which configure the op amp for a *finite* gain *K*, multiple-feedback filters exploit the full open-loop gain and are also referred to as *infinite-gain filters*. Together with *KRC* filters, they are the most popular single-op-amp realizations of the second-order responses.



(*a*)



(*b*)

FIGURE 3.29
(*a*) PSpice circuit for the notch filter of Example 3.14, and (*b*) its magnitude response.

Band-Pass Filters

In the circuit of Fig. 3.30, also called the *Delyiannis-Friend* filter, named after its inventors, the op amp acts as a differentiator with respect to V_1, so we write

$$V_o = -sR_2C_2V_1$$

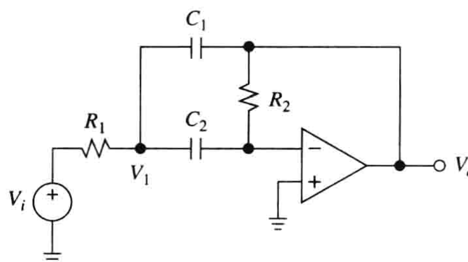


FIGURE 3.30
Multiple-feedback band-pass filter.

Summing currents at node V_1,

$$\frac{V_i - V_1}{R_1} + \frac{V_o - V_1}{1/sC_1} + \frac{0 - V_1}{1/sC_2} = 0$$

Eliminating V_1, letting $s \to j\omega$, and rearranging,

$$H(j\omega) = \frac{V_o}{V_i} = \frac{-j\omega R_2 C_2}{1 - \omega^2 R_1 R_2 C_1 C_2 + j\omega R_1(C_1 + C_2)}$$

To put this function in the standard form $H(j\omega) = H_{0BP} H_{BP}(j\omega)$, we impose $\omega^2 R_1 R_2 C_1 C_2 = (\omega/\omega_0)^2$ to get

$$\omega_0 = \frac{1}{\sqrt{R_1 R_2 C_1 C_2}} \tag{3.70a}$$

and $j\omega R_1(C_1 + C_2) = (j\omega/\omega_0)/Q$ to get

$$Q = \frac{\sqrt{R_2/R_1}}{\sqrt{C_1/C_2} + \sqrt{C_2/C_1}} \tag{3.70b}$$

Finally, we impose $-j\omega R_2 C_2 = H_{0BP} \times (j\omega/\omega_0)/Q$ to get

$$H_{0BP} = \frac{-R_2/R_1}{1 + C_1/C_2} \tag{3.70c}$$

Clearly, this filter is of the inverting type. It is customary to impose $C_1 = C_2 = C$, after which the above expressions simplify to

$$\omega_0 = \frac{1}{\sqrt{R_1 R_2} C} \qquad Q = 0.5\sqrt{R_2/R_1} \qquad H_{0BP} = -2Q^2 \tag{3.71}$$

The corresponding design equations are

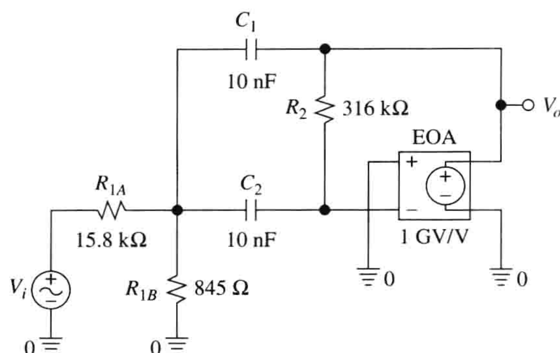
$$R_1 = 1/2\omega_0 QC \qquad R_2 = 2Q/\omega_0 C \tag{3.72}$$

Denoting resonance-gain magnitude as $H_0 = |H_{0BP}|$ for simplicity, we observe that it increases quadratically with Q. If we want $H_0 < 2Q^2$, we must replace R_1 with a voltage divider in the manner of Example 3.9. The design equations are then

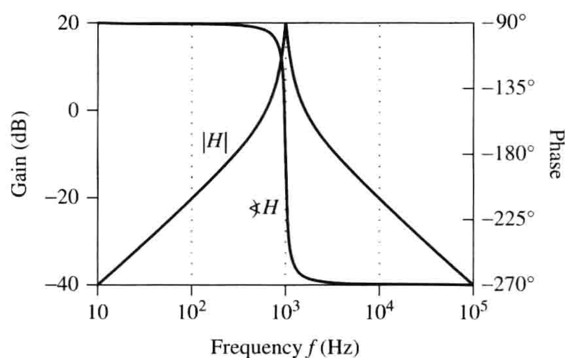
$$R_{1A} = Q/H_0\omega_0 C \qquad R_{1B} = R_{1A}/(2Q^2/H_0 - 1) \tag{3.73}$$

EXAMPLE 3.15. Design a multiple-feedback band-pass filter with $f_0 = 1$ kHz, $Q = 10$, and $H_0 = 20$ dB. Check with PSpice.

Solution. Let $C_1 = C_2 = 10$ nF. Then, $R_2 = 2 \times 10/(2\pi 10^3 \times 10^{-8}) = 318.3$ kΩ (use 316 kΩ, 1%). Since 20 dB implies $H_0 = 10$ V/V, which is less than $2Q^2 = 200$, we need an input attenuator. Thus, $R_{1A} = 10/(10 \times 2\pi 10^3 \times 10^{-8}) = 15.92$ kΩ (use 15.8 kΩ, 1%), and $R_{1B} = 15.92/(200/10 - 1) = 837.7$ Ω (use 845 Ω, 1%). Using the circuit of Fig. 3.31*a*, we get the magnitude plot of Fig. 3.31*b*. Cursor measurements give $|H_{0BP}| = 20.0$ dB, $f_0 = 999.7$ Hz, and $20 - 3 = 17$-dB frequencies of 950.6 Hz and 1051.3 Hz, so $Q = 999.7/(1051.3 - 950.6) = 10.0$. (It is interesting to compare the phase plot with that of Fig. 3.25*b*. Can you justify the difference?)



(a)



(b)

FIGURE 3.31
(a) PSpice circuit for the band-pass filter of Example 3.15, and (b) its Bode plots.

Low-Pass Filters

The circuit of Fig. 3.32 consists of the low-pass stage R_1-C_1 followed by the integrator stage made up of R_2, C_2, and the op amp, so we anticipate a low-pass response. Moreover, the presence of positive feedback via R_3 should allow for Q control. The

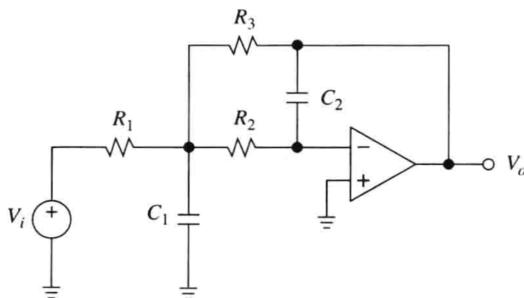


FIGURE 3.32
Multiple-feedback low-pass filter.

ac analysis of the circuit gives $V_o/V_i = H_{0LP}H_{LP}$, where

$$H_{0LP} = -\frac{R_3}{R_1} \qquad \omega_0 = \frac{1}{\sqrt{R_2R_3C_1C_2}}$$
$$Q = \frac{\sqrt{C_1/C_2}}{\sqrt{R_2R_3/R_1^2} + \sqrt{R_3/R_2} + \sqrt{R_2/R_3}} \tag{3.74}$$

These expressions indicate that we can vary R_3 to adjust ω_0, and R_1 to adjust Q.

EXERCISE 3.4. Derive Eq. (3.74).

A possible design procedure[2] is to choose a convenient value for C_2 and calculate $C_1 = nC_2$, where n is the capacitance spread,

$$n \geq 4Q^2(1 + H_0) \tag{3.75}$$

H_0 being the desired dc-gain magnitude. The resistances are then found as

$$R_3 = \frac{1 + \sqrt{1 - 4Q^2(1 + H_0)/n}}{2\omega_0 QC_2} \qquad R_1 = \frac{R_3}{H_0} \qquad R_2 = \frac{1}{\omega_0^2 R_3C_1C_2} \tag{3.76}$$

A disadvantage of this filter is that the higher the Q and H_0, the greater the capacitance spread.

EXAMPLE 3.16. Design a multiple-feedback low-pass filter with $H_0 = 2$ V/V, $f_0 = 10$ kHz, and $Q = 4$.

Solution. Substituting the given values yields $n \geq 192$. Let $n = 200$. Start out with $C_2 = 1$ nF. Then, $C_1 = 0.2\ \mu$F, $R_3 = 2.387$ kΩ (use 2.37 kΩ, 1%), $R_1 = 1.194$ kΩ (use 1.18 kΩ, 1%), and $R_3 = 530.5$ Ω (use 536 Ω, 1%).

Notch Filters

The circuit of Fig. 3.33 exploits Eq. (3.55) to synthesize the notch response using the band-pass response. By inspection, $V_o = -(R_5/R_3)(-H_0H_{BP})V_i - (R_5/R_4)V_i = -(R_5/R_4)[1 - (H_0R_4/R_3)H_{BP}]V_i$. It is apparent that imposing $H_0R_4/R_3 = 1$ leads to a mutual cancellation of the $(j\omega/\omega_0)/Q$ terms in the numerator, giving $V_o/V_i = H_{0N}H_N$, $H_{0N} = -R_5/R_4$.

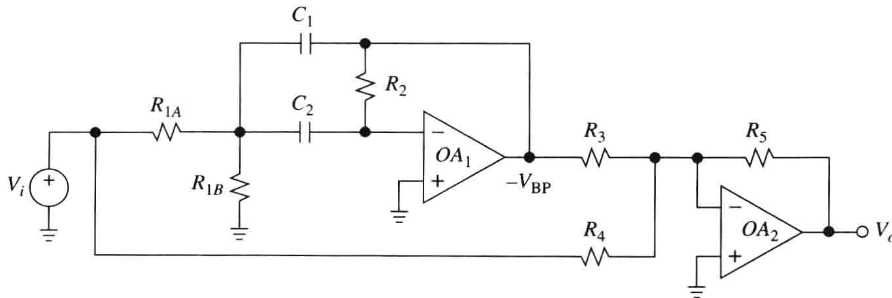


FIGURE 3.33
Synthesis of H_N using H_{BP}.

EXAMPLE 3.17. Design a notch filter with $f_0 = 1$ kHz, $Q = 10$, and $H_{0N} = 0$ dB.

Solution. First, implement a band-pass stage with $f_0 = 1$ kHz, $Q = 10$, and $H_0 = 1$ V/V. Using $C_1 = C_2 = 10$ nF, this requires $R_2 = 318.3$ kΩ, $R_{1A} = 159.2$ kΩ, and $R_{1B} = 799.8$ Ω. Then, use $R_3 = R_4 = R_5 = 10.00$ kΩ.

3.7 STATE-VARIABLE AND BIQUAD FILTERS

The second-order filters investigated so far use a single op amp with a minimum or near-minimum number of external components. Simplicity, however, does not come without a price. Drawbacks such as wide component spreads, awkward tuning capabilities, and high sensitivity to component variations, particularly to the gain of the amplifier, generally limit these filters to $Q \leq 10$.

Component minimization, especially minimization of the number of op amps, was of concern when these devices were expensive. Nowadays, multiple-op-amp packages such as duals and quads are cost-competitive with precision passive components. The question then arises whether filter performance and versatility can be improved by shifting the burden from passive to active devices. The answer is provided by multiple-op-amp filters, such as the *state-variable* and *biquad types,* which, though using more components, are generally easier to tune, are less sensitive to passive component variations, and do not require extravagant component spreads. Since they provide more than one response simultaneously, they are also referred to as *universal filters.*

State-Variable (SV) Filters

The SV filter—also known as the *KHN filter* for inventors W. J. Kerwin, L. P. Huelsman, and R. W. Newcomb, who first reported it in 1967—uses two integrators and a summing amplifier to provide the second-order low-pass, band-pass, and high-pass responses. A fourth op amp can be used to combine the existing responses and synthesize the notch or the all-pass responses. The circuit realizes a second-order differential equation, hence its name.

In the SV version of Fig. 3.34, OA_1 forms a linear combination of the input and the outputs of the remaining op amps. Using the superposition principle, we write

$$V_{\mathrm{HP}} = -\frac{R_5}{R_3}V_i - \frac{R_5}{R_4}V_{\mathrm{LP}} + \left(1 + \frac{R_5}{R_3 \parallel R_4}\right)\frac{R_1}{R_1 + R_2}V_{\mathrm{BP}}$$

$$= -\frac{R_5}{R_3}V_i - \frac{R_5}{R_4}V_{\mathrm{LP}} + \frac{1 + R_5/R_3 + R_5/R_4}{1 + R_2/R_1}V_{\mathrm{BP}} \tag{3.77}$$

Since OA_2 and OA_3 are integrators, we have

$$V_{\mathrm{BP}} = \frac{-1}{R_6 C_1 s}V_{\mathrm{HP}} \qquad V_{\mathrm{LP}} = \frac{-1}{R_7 C_2 s}V_{\mathrm{BP}} \tag{3.78}$$

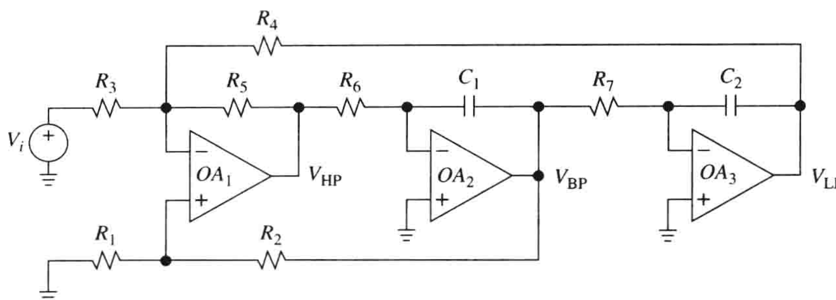


FIGURE 3.34
State-variable filter (inverting).

or $V_{LP} = (1/R_6C_1R_7C_2s^2)V_{HP}$. Substituting V_{BP} and V_{LP} into Eq. (3.77) and collecting, we get

$$\frac{V_{HP}}{V_i} = -\frac{R_5}{R_3}\frac{\frac{R_4R_6C_1R_7C_2}{R_5}s^2}{\frac{R_4R_6C_1R_7C_2}{R_5}s^2 + \frac{R_4R_7C_2(1 + R_5/R_3 + R_5/R_4)}{R_5(1 + R_2/R_1)}s + 1}$$

Putting this expression in the standard form $V_{HP}/V_i = H_{0HP}H_{HP}$ allows us to find $H_{0HP} = -R_5/R_3$ and

$$\omega_0 = \frac{\sqrt{R_5/R_4}}{\sqrt{R_6C_1R_7C_2}} \qquad Q = \frac{(1 + R_2/R_1)\sqrt{R_5R_6C_1/R_4R_7C_2}}{1 + R_5/R_3 + R_5/R_4} \tag{3.79}$$

Using $V_{BP}/V_i = (-1/R_6C_1s)V_{HP}/V_i$ indicates that $V_{BP}/V_i = H_{0BP}H_{BP}$ and also allows us to find H_{0BP}. We similarly find $V_{LP}/V_i = (-1/R_7C_2s)V_{BP}/V_i = H_{0LP}H_{LP}$. The results are

$$H_{0HP} = -\frac{R_5}{R_3} \qquad H_{0BP} = \frac{1 + R_2/R_1}{1 + R_3/R_4 + R_3/R_5} \qquad H_{0LP} = -\frac{R_4}{R_3} \tag{3.80}$$

The above derivations reveal some interesting properties: first, the band-pass response is generated by integrating the high-pass response, and the low-pass is in turn generated by integrating the band-pass; second, since the product of two transfer functions corresponds to the addition of their decibel plots, and since the integrator plot has a constant slope of −20 dB/dec, the band-pass decibel plot is obtained by rotating the high-pass decibel plot clockwise by 20 dB/dec and the low-pass plot by a similar rotation of the band-pass plot.

We observe that Q is no longer the result of a cancellation, as in the case of *KRC* filters, but depends on the resistor ratio R_2/R_1 in a straightforward manner. We therefore expect Q to be much less sensitive to resistance tolerances and drift. Indeed, with proper component selection and circuit construction, the SV filter can easily yield dependable Qs in the range of hundreds. For best results, use metal-film resistors and polystyrene or polycarbonate capacitors, and properly bypass the op amp supplies.

The SV filter is usually implemented with $R_5 = R_4 = R_3$, $R_6 = R_7 = R$, and $C_1 = C_2 = C$, so the earlier expressions simplify to

$$\omega_0 = 1/RC \qquad Q = \frac{1}{3}(1 + R_2/R_1) \tag{3.81a}$$

$$H_{0\text{HP}} = -1 \qquad H_{0\text{BP}} = Q \qquad H_{0\text{LP}} = -1 \tag{3.81b}$$

The filter is tuned as follows: (*a*) adjust R_3 for the desired magnitude of the response of interest; (*b*) adjust R_6 (or R_7) to tune ω_0; (*c*) adjust the ratio R_2/R_1 to tune Q.

EXAMPLE 3.18. In the circuit of Fig 3.34 specify component values for a band-pass response with a bandwidth of 10 Hz centered at 1 kHz. What is the resonance gain?

Solution. Pick the convenient values $C_1 = C_2 = 10$ nF. Then, $R = 1/(2\pi 10^3 \times 10^{-8}) = 15.92$ kΩ (use 15.8 kΩ, 1%). By definition, $Q = f_0/\text{BW} = 10^3/10 = 100$. Imposing $(1 + R_2/R_1)/3 = 100$ gives $R_2/R_1 = 299$. Pick $R_1 = 1.00$ kΩ, 1%, and $R_2 = 301$ kΩ, 1%. To simplify inventory, let also $R_3 = R_4 = R_5 = 15.8$ kΩ, 1%. The gain at resonance is $H_{0\text{BP}} = 100$ V/V.

Equation (3.81*b*) indicates that at $\omega = \omega_0$ all three responses exhibit a magnitude of Q V/V. In high-Q situations this may cause the op amps to saturate, unless the input signal level is kept suitably low. Low-input levels can be obtained by replacing R_3 with a suitable voltage divider, in the manner of Example 3.9 (see Problem 3.42).

Moving the input signal from the inverting to the noninverting side of OA_1 results in the circuit of Fig. 3.35, which represents another popular form of the SV filter. It can be proved (see Problem 3.43) that with the components shown, we now have

$$\omega_0 = 1/RC \qquad Q = 1 + R_2/2R_1 \tag{3.82a}$$

$$H_{0\text{HP}} = 1/Q \qquad H_{0\text{BP}} = -1 \qquad H_{0\text{LP}} = 1/Q \tag{3.82b}$$

indicating that for $\omega = \omega_0$ all three responses now exhibit 0-dB magnitudes. The band-pass plot is as in Fig. 3.20*a*; the low- and high-pass plots are as in Fig. 3.19, but shifted downward by Q_{dB}.

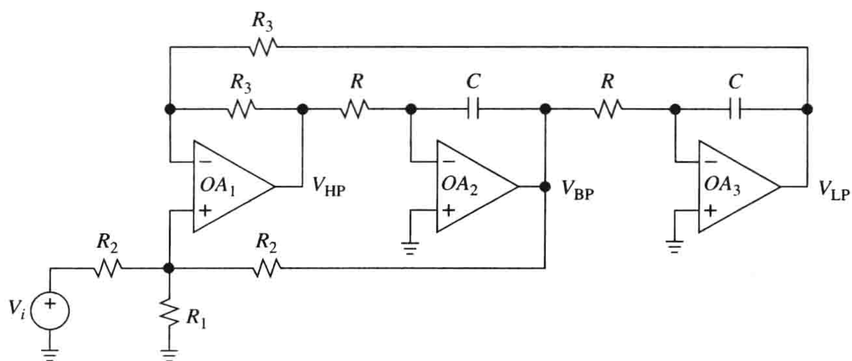


FIGURE 3.35
State-variable filter (noninverting).

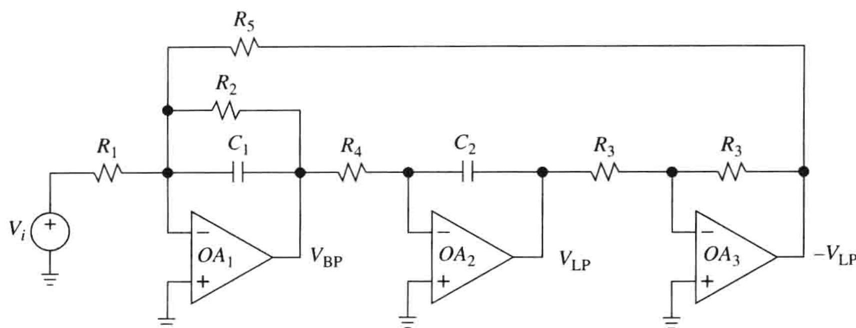


FIGURE 3.36
Biquad filter.

The Biquad Filter

Also known as the *Tow-Thomas filter,* for its inventors, the circuit of Fig. 3.36 consists of two integrators, one of which is of the lossy type. The third op amp is a unity-gain inverting amplifier whose sole purpose is to provide polarity reversal. If one of the integrators is allowed to be of the noninverting type, the inverting amplifier is omitted and only two op amps are required.

To analyze the circuit, we sum currents at the inverting-input node of OA_1,

$$\frac{V_i}{R_1} + \frac{-V_{\text{LP}}}{R_5} + \frac{V_{\text{BP}}}{R_2} + \frac{V_{\text{BP}}}{1/sC_1} = 0$$

Letting $V_{\text{LP}} = (-1/R_4C_2s)V_{\text{BP}}$ and collecting gives $V_{\text{BP}}/V_i = H_{0\text{BP}}H_{\text{BP}}$ and $V_{\text{LP}}/V_i = (-1/R_4C_2s)V_{\text{BP}}/V_i = H_{0\text{LP}}H_{\text{LP}}$, with

$$H_{0\text{BP}} = -\frac{R_2}{R_1} \qquad H_{0\text{LP}} = \frac{R_5}{R_1} \qquad \omega_0 = \frac{1}{\sqrt{R_4R_5C_1C_2}} \qquad Q = \frac{R_2\sqrt{C_1}}{\sqrt{R_4R_5C_2}} \tag{3.83}$$

We observe that unlike the SV filter, the biquad yields only two significant responses. However, since all its op amps are operated in the inverting mode, the circuit is immune from common-mode limitations, an issue to be studied in Chapter 5.

The biquad filter is usually implemented with $R_4 = R_5 = R$ and $C_1 = C_2 = C$, after which the above expressions simplify as

$$H_{0\text{BP}} = -\frac{R_2}{R_1} \qquad H_{0\text{LP}} = \frac{R}{R_1} \qquad \omega_0 = \frac{1}{RC} \qquad Q = \frac{R_2}{R} \tag{3.84}$$

The filter is tuned as follows: (*a*) adjust R_4 (or R_5) to tune ω_0; (*b*) adjust R_2 to tune Q; (*c*) adjust R_1 for the desired value of $H_{0\text{BP}}$ or of $H_{0\text{LP}}$.

EXAMPLE 3.19. Design a biquad filter with $f_0 = 8$ kHz, BW = 200 Hz, and a 20-dB resonance gain. What is the value of $H_{0\text{LP}}$?

Solution. Let $C_1 = C_2 = 1$ nF. Then, $R_4 = R_5 = 1/(2\pi \times 8 \times 10^3 \times 10^{-9}) = 19.89$ kΩ (use 20.0 kΩ, 1%); $Q = 8 \times 10^3/200 = 40$; $R_2 = 40 \times 19.89 = 795.8$ kΩ (use 787 kΩ, 1%); $R_1 = R_2/10^{20/20} = 78.7$ kΩ, 1%; $H_{0\text{LP}} = 20.0/78.7 = 0.254$ V/V, or -11.9 dB.

The Notch Response

With the help of a fourth op amp and a few resistors, both the biquad and the SV circuits can be configured for the notch response, which explains why these filters are also called *universal.* With a quad package, the fourth op amp is already available, so it only takes a few resistors to synthesize a notch.

The filter of Fig. 3.37 uses the biquad circuit to generate the notch response as $V_N = -[(R_5/R_2)(V_i - V_{BP}) \pm (R_5/R_4)V_{LP}]$, where the $\pm$ sign depends on the switch position, as indicated. It can be shown (see Exercise 3.5) that

$$\frac{V_N}{V_i} = -\frac{R_5\omega_z^2}{R_2\omega_0^2} \times \frac{1-(\omega/\omega_z)^2}{1-(\omega/\omega_0)^2 + (j\omega/\omega_0)/Q} \tag{3.85a}$$

$$\omega_0 = \frac{1}{RC} \qquad Q = \frac{R_1}{R} \qquad \omega_z = \omega_0\sqrt{1 \pm R_2/R_4Q} \tag{3.85b}$$

This response presents a notch at $\omega = \omega_z$. We identify three cases:

1. R_4 is absent, or $R_4 = \infty$. By Eq. (3.85), we have

$$\omega_z = \omega_0 \qquad H_{0N} = -\frac{R_5}{R_2} \tag{3.86}$$

   This is the familiar *symmetric notch* shown in Fig. 3.38*b* for the case $|H_{0N}| = 0$ dB. It is obtained by subtracting V_{BP} from V_i, in the manner depicted in Fig. 3.33.
2. The switch is in the left position, so also a low-pass term is now being added to the existing combination of V_i and $-V_{BP}$. The result is a *low-pass notch.* By

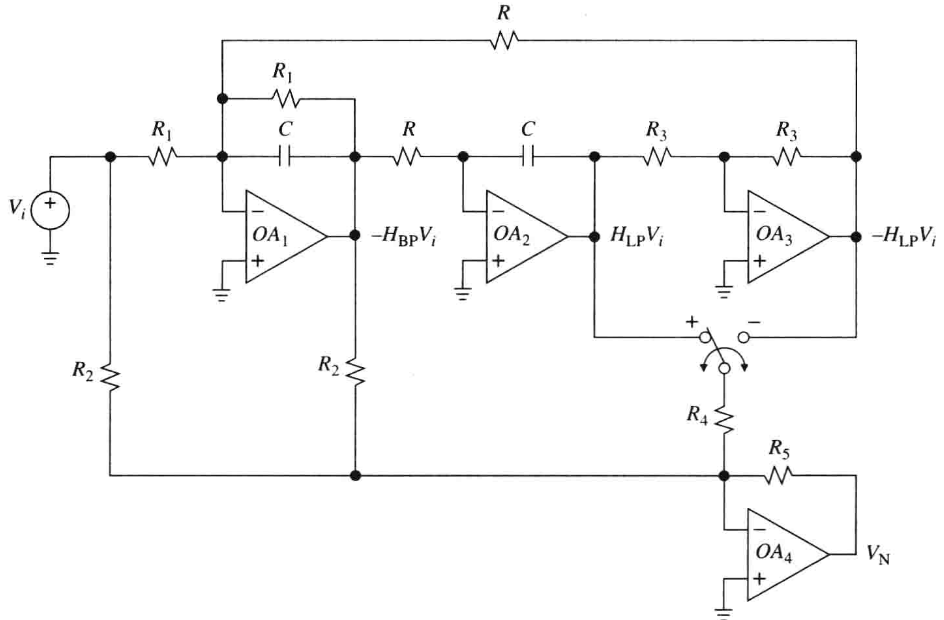


FIGURE 3.37
Synthesizing notch responses.

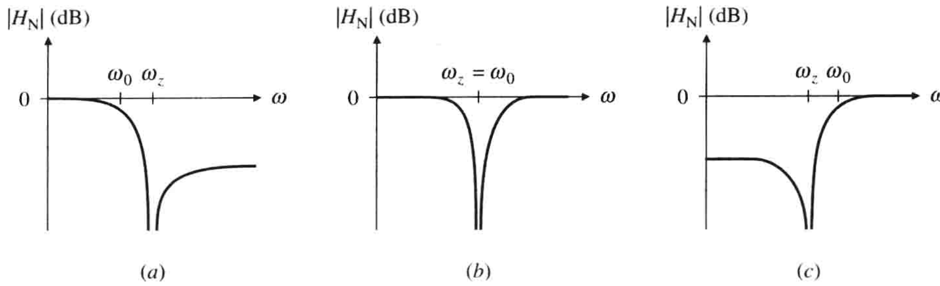


FIGURE 3.38
Notch responses: (*a*) low-pass notch, (*b*) symmetric notch, and (*c*) high-pass notch.

Eq. (3.85), we now have

$$\omega_z = \omega_0\sqrt{1 + R_2/R_4Q} \qquad H_{0\mathrm{LP}} = -\frac{R_5\omega_z^2}{R_2\omega_0^2} \tag{3.87}$$

indicating $\omega_z > \omega_0$. The scaling term is called the *dc gain* $H_{0\mathrm{LP}}$. The low-pass notch is shown in Fig. 3.38*a* for the case $|H_{0\mathrm{LP}}| = 0$ dB. By Eq. (3.85*a*), the high-frequency gain is $H_{0\mathrm{HP}} = H_{0\mathrm{LP}}(1/\omega_z^2)/(1/\omega_0^2) = -R_5/R_2$.

3. The switch is in the right position, so the low-pass term is now being subtracted. The result is a *high-pass notch* with

$$\omega_z = \omega_0\sqrt{1 - R_2/R_4Q} \qquad H_{0\mathrm{HP}} = -\frac{R_5}{R_2} \tag{3.88}$$

We now have $\omega_z < \omega_0$, and the scaling factor is called the *high-frequency gain* $H_{0\mathrm{HP}}$. This notch is shown in Fig. 3.38*c* for the case $|H_{0\mathrm{HP}}| = 0$ dB. The dc gain is $H_{0\mathrm{LP}} = -R_5\omega_z^2/R_2\omega_0^2$.

EXERCISE 3.5. Derive Eq. (3.85).

In Chapter 4 we shall use low- and high-pass notches to synthesize a class of higher-order filters known as *elliptic filters.* The above expressions can be turned around to yield the design equations:

$$R = \frac{1}{\omega_0 C} \qquad R_1 = QR \qquad R_4 = \frac{R_2}{Q}\frac{\omega_0^2}{|\omega_0^2 - \omega_z^2|} \tag{3.89a}$$

$$R_5 = R_2\left(\frac{\omega_0}{\omega_z}\right)^2 \quad \text{for } \omega_z > \omega_0 \qquad R_5 = R_2 \quad \text{for } \omega_z < \omega_0 \tag{3.89b}$$

where R_2 and R_3 are arbitrary and R_5 has been specified for $H_{0\mathrm{LP}}$ and $H_{0\mathrm{HP}}$ of 0 dB. These gains can be raised or lowered by changing R_5 in proportion.

EXAMPLE 3.20. Specify the components of Fig. 3.37 for a low-pass notch with $f_0 =$ 1 kHz, $f_z = 2$ kHz, $Q = 10$, and 0-dB dc gain. Verify with PSpice.

Solution. Let $C = 10$ nF; then $R = 1/\omega_0 C = 15.9$ kΩ (use 15.8 kΩ); $R_1 = QR =$ 158 kΩ; let $R_2 = 100$ kΩ; then $R_4 = (100/10) \times 1^2/|1^2 - 2^2| = 3.333$ kΩ (use 3.32 kΩ, 1%); $R_5 = 100 \times (1/2)^2 = 25$ kΩ (use 24.9 kΩ, 1%). Using the PSpice circuit of Fig. 3.39 (top) we get the magnitude plot of Fig. 3.39 (bottom). Cursor measurements give $f_z = 2.0177$ kHz and a notch depth of -105 dB.

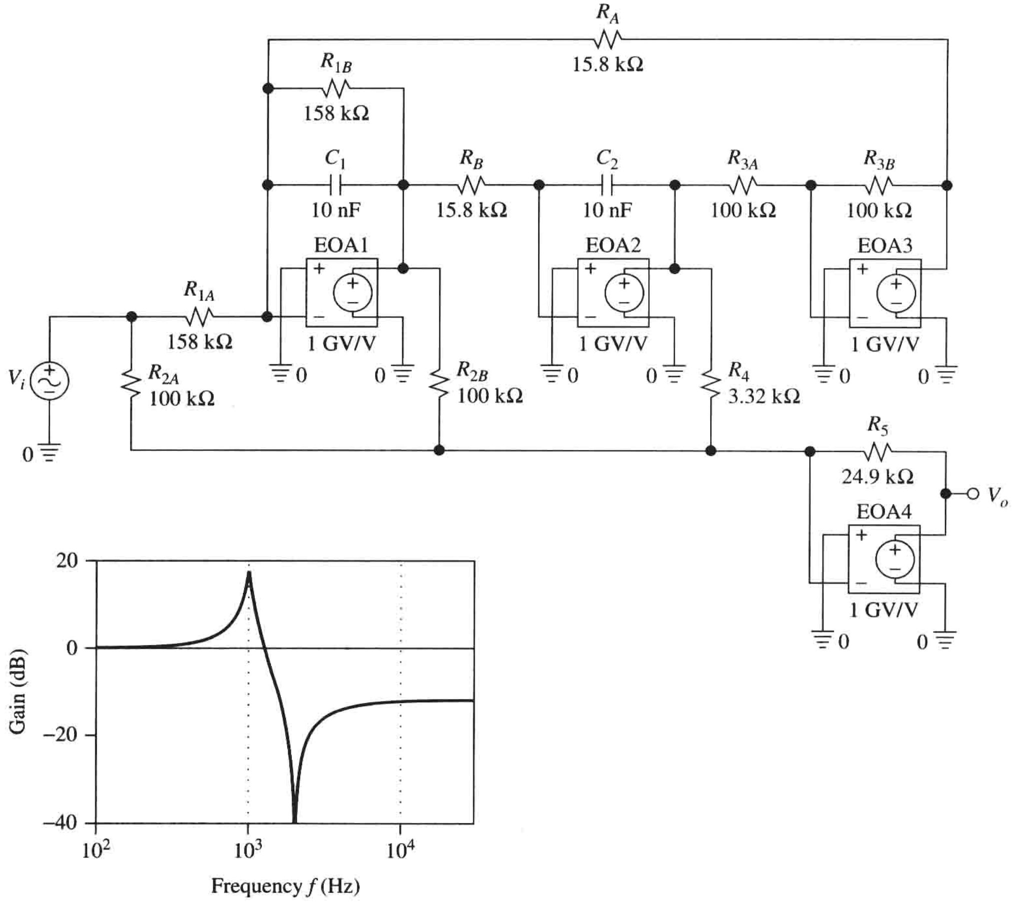


FIGURE 3.39
PSpice circuit for the low-pass notch filter of Example 3.20, and its frequency response.

3.8 SENSITIVITY

Because of component tolerances and op amp nonidealities, the response of a practical filter is likely to deviate from that predicted by theory. Even if some of the components are made adjustable to allow for fine tuning, deviations will still arise because of component aging and thermal drift. It is therefore of interest to know how sensitive a given filter is to component variations. For instance, the designer of a second-order band-pass filter may want to know the extent to which a 1% variation in a given resistance or capacitance affects ω_0 and BW.

Given a filter parameter y such as ω_0 and Q, and given a filter component x such as a resistance R or a capacitance C, the *classical sensitivity function* S_x^y is defined as

$$S_x^y = \frac{\partial y/y}{\partial x/x} = \frac{x}{y}\frac{\partial y}{\partial x} \tag{3.90}$$

where we use partial derivatives to account for the fact that filter parameters usually depend on more than just one component. For small changes, we can approximate

$$\frac{\Delta y}{y} \cong S_x^y \frac{\Delta x}{x} \qquad (3.91)$$

This allows us to estimate the *fractional parameter change* $\Delta y/y$ caused by the *fractional component change* $\Delta x/x$. Multiplying both sides by 100 gives a relationship between *percentage changes*. The sensitivity function satisfies the following useful properties:

$$S_{1/x}^y = S_x^{1/y} = -S_x^y \qquad (3.92a)$$

$$S_x^{y_1 y_2} = S_x^{y_1} + S_x^{y_2} \qquad (3.92b)$$

$$S_x^{y_1/y_2} = S_x^{y_1} - S_x^{y_2} \qquad (3.92c)$$

$$S_x^{x^n} = n \qquad (3.92d)$$

$$S_{x_1}^y = S_{x_2}^y S_{x_1}^{x_2} \qquad (3.92e)$$

(See Problem 3.49 for the derivations.) To gain an understanding of sensitivity, we examine some popular filter configurations.

KRC Filter Sensitivities

With reference to the *low-pass KRC filter* of Fig. 3.23, we have, by Eq. (3.60*b*), $\omega_0 = R_1^{-1/2} C_1^{-1/2} R_2^{-1/2} C_2^{-1/2}$. Consequently, Eq. (3.92*d*) gives

$$S_{R_1}^{\omega_0} = S_{C_1}^{\omega_0} = S_{R_2}^{\omega_0} = S_{C_2}^{\omega_0} = -\frac{1}{2} \qquad (3.93)$$

Applying Eqs. (3.90) and (3.92) to the expression for Q given in Eq. (3.60*c*), we obtain

$$S_{R_1}^Q = -S_{R_2}^Q = Q\sqrt{R_2C_2/R_1C_1} - \frac{1}{2} \qquad (3.94a)$$

$$S_{C_1}^Q = -S_{C_2}^Q = Q(\sqrt{R_2C_2/R_1C_1}) + \sqrt{R_1C_2/R_2C_1}) - \frac{1}{2} \qquad (3.94b)$$

$$S_K^Q = QK\sqrt{R_1C_1/R_2C_2} \qquad (3.94c)$$

$$S_{R_A}^Q = -S_{R_B}^Q = Q(1-K)\sqrt{R_1C_1/R_2C_2} \qquad (3.94d)$$

For the *equal-component* design, the Q sensitivities simplify to

$$S_{R_1}^Q = -S_{R_2}^Q = Q - \frac{1}{2} \qquad S_{C_1}^Q = -S_{C_2}^Q = 2Q - \frac{1}{2} \qquad (3.95a)$$

$$S_K^Q = 3Q - 1 \qquad S_{R_A}^Q = -S_{R_B}^Q = 1 - 2Q \qquad (3.95b)$$

and for the *unity-gain* design, they simplify to

$$S_{R_1}^Q = -S_{R_2}^Q = \frac{1 - R_1/R_2}{2(1 + R_1/R_2)} \qquad S_{C_1}^Q = -S_{C_2}^Q = \frac{1}{2} \qquad (3.96)$$

Since the Q sensitivities of the equal-component design increase with Q, they may become unacceptable at high Qs. As we already know, S_K^Q is of particular concern at high Qs because a slight mismatch in the R_B/R_A ratio may drive Q to infinity or even make it negative, thus leading to oscillatory behavior. By contrast, the unity-gain design offers much lower sensitivities. It is apparent that the designer must carefully weigh a number of conflicting factors before choosing a particular filter design for the given application. These include circuit simplicity, cost, component spread, tunability, and sensitivity.

EXAMPLE 3.21. Investigate the effect of a 1% variation of each component in the low-pass filter of (*a*) Example 3.8 and (*b*) Example 3.10.

Solution. By Eq. (3.93), a 1% increase (decrease) in any of R_1, C_1, R_2, and C_2 causes a 0.5% decrease (increase) in ω_0 in either circuit.

(*a*) By Eq. (3.95), a 1% increase (decrease) in R_1 increases (decreases) Q by approximately $5 - 0.5 = 4.5\%$ (the opposite holds for R_2). Similarly, 1% capacitance variations result in Q variations of about 9.5%. Finally, since $1 - 2Q = 1 - 2 \times 5 = -9$, it follows that 1% variations in R_A or in R_B result in Q variations of about 9%.

(*b*) With $R_1/R_2 = 5.76/2.21$, Eq. (3.96) gives $S_{R_1}^Q = -S_{R_2}^Q \cong -0.22$. Thus, 1% resistance and 1% capacitance variations result in Q variations of 0.22% and 0.5%, respectively.

Multiple-Feedback Filter Sensitivities

The sensitivities of the *multiple-feedback band-pass filter* of Fig. 3.30 are found from Eq. (3.70), and they are

$$S_{R_1}^{\omega_0} = S_{C_1}^{\omega_0} = S_{R_2}^{\omega_0} = S_{C_2}^{\omega_0} = -\frac{1}{2} \qquad (3.97a)$$

$$S_{R_1}^{Q} = -S_{R_2}^{Q} = -\frac{1}{2} \qquad S_{C_1}^{Q} = -S_{C_2}^{Q} = \frac{1}{2}\frac{C_2 - C_1}{C_2 + C_1} \qquad (3.97b)$$

Note that the equal-capacitance design results in $S_{C_1}^Q = S_{C_2}^Q = 0$. The sensitivities of the *multiple-feedback low-pass filter* of Fig. 3.32 can be computed likewise, and they are found to be[2]

$$S_{R_2}^{\omega_0} = S_{C_1}^{\omega_0} = S_{R_3}^{\omega_0} = S_{C_2}^{\omega_0} = -\frac{1}{2} \qquad (3.98a)$$

$$|S_{R_1}^{Q}| < 1 \qquad |S_{R_2}^{Q}| < \frac{1}{2} \qquad |S_{R_3}^{Q}| < \frac{1}{2} \qquad S_{C_1}^{Q} = -S_{C_2}^{Q} = \frac{1}{2} \qquad (3.98b)$$

It is apparent that multiple-feedback configurations enjoy low sensitivities and are therefore popular.

Multiple-Op-Amp Filter Sensitivities

The sensitivities of the *biquad filter* of Fig. 3.36 are found from Eq. (3.83), and the results are

$$S_{R_4}^{\omega_0} = S_{R_5}^{\omega_0} = S_{C_1}^{\omega_0} = S_{C_2}^{\omega_0} = -\frac{1}{2} \qquad (3.99a)$$

$$S_{R_2}^{Q} = 1 \qquad S_{R_4}^{Q} = S_{R_5}^{Q} = -S_{C_1}^{Q} = S_{C_2}^{Q} = \frac{1}{2} \qquad (3.99b)$$

These sensitivities are fairly low and are similar to those of a passive *RLC* filter yielding the same responses. The sensitivities of *state-variable filters* are similarly low (see Problem 3.52). Considering also the advantages of tuning, low parameter spread, and multiple simultaneous responses, we now appreciate why these filters are widely used.

PROBLEMS

3.1 The transfer function

3.1 Use PSpice to plot the impulse response of Example 3.2, compare with the calculated response, and comment. *Hint:* You can approximate an impulse by means of a pulse having much shorter width than the lowest time constant of the circuit and sufficient amplitude to ensure unity area. Presently, a 1-μs, 1 MV pulse will do. Also, make sure to impose zero initial conditions for L and C.

3.2 A transfer function with $H_0 = 1$ has a zero at $s = +1$ kNp/s and a pole pair at $-1 \pm j1$ complex kNp/s. (a) Find its impulse response. (b) Find its steady-state response to an ac input with unity amplitude, zero phase, and $\omega = 1$ krad/s.

3.3 Suppose we lower R to 4 Ω in the circuit of Fig. 3.3a, and we take the output $v_o(t)$ across C (positive at left). (a) Use PSpice with zero initial conditions for L and C to display the response $v_o(t)$ to a 1-V input step. How long does it take for the response to settle reasonably close to its final dc steady-state value? (b) Now display the response to a 10-V, 2-kHz sinusoidal input. How long does it take for the response to settle reasonably close to its final ac steady-state? Compare with (a), and comment.

3.2 First-order active filters

3.4 The circuit of Fig. P3.4 is a noninverting differentiator. (a) Derive its transfer function. (b) Specify component values for a unity-gain frequency of 100 Hz.

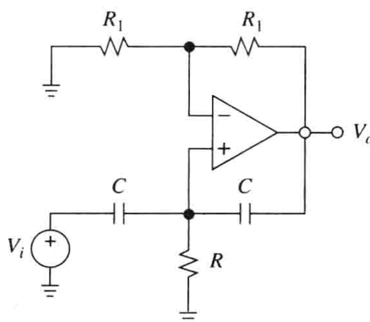


FIGURE P3.4

3.5 If $R_1C_1 = R_2C_2$, the circuit of Fig. P3.5 is a noninverting integrator. (a) Find its transfer function. (b) Specify component values for a gain of 20 dB at 100 Hz.

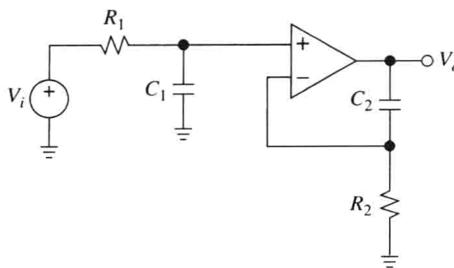


FIGURE P3.5

3.6 (*a*) Specify suitable component values for a unity-gain frequency of 1 kHz in the Deboo integrator of Fig. 3.7. (*b*) What happens if the upper-right resistance is 1% less than its nominal value? Illustrate via the magnitude plot. *Hint:* Replace the Howland current pump with its Norton equivalent.

3.7 Suppose the time constants in the circuit of Fig. P3.5 are mismatched, say, $R_1C_1 = R_2C_2(1-\epsilon)$. (*a*) Investigate the effect of the mismatch and illustrate via the magnitude plot. (*b*) Devise a method for balancing out the mismatch, and outline the calibration procedure.

3.8 Inserting a resistance R_3 in series with C in the low-pass filter of Fig. 3.9*a* turns it into a circuit known as a *pole-zero* circuit, which finds application in control. (*a*) Sketch the modified circuit, and find its transfer function to justify its name. (*b*) Specify component values for a pole frequency of 1 kHz, a zero frequency of 10 kHz, and a dc gain of 0 dB; sketch its magnitude plot.

3.9 Inserting a resistance R_3 in parallel with C in the high-pass filter of Fig 3.10*a* turns it into a circuit known as a *zero-pole* circuit, which finds application in control. (*a*) Sketch the modified circuit, and find its transfer function to justify its name. (*b*) Specify component values for a zero frequency of 100 Hz, a pole frequency of 1 kHz, and a high-frequency gain of 0 dB; sketch its magnitude plot.

3.10 Redraw the phase shifter of Fig 3.12*a*, but with R and C interchanged with each other; derive its transfer function and sketch its Bode plots. What is the main difference between the responses of the original and the modified circuit? Name a possible disadvantage of the modified circuit.

3.11 (*a*) Sketch the Bode plots of the circuit of Fig 3.12*a* if $R_2 = 10R_1$. (*b*) Repeat, but with $R_1 = 10R_2$.

3.12 Using two phase shifters with 0.1-μF capacitors, design a circuit that accepts a voltage $v_a = 1.20\sqrt{2}\cos(2\pi 60t)$ V, and generates the voltages $v_b = 1.20\sqrt{2}\cos(2\pi 60t - 120°)$ V and $v_c = 1.20\sqrt{2}\cos(2\pi 60t + 120°)$ V. Such a circuit simulates the voltages used in three-phase power transmission systems, scaled to 1/100 of their actual values.

3.13 In the noninverting amplifier of Fig. 1.7 let $R_1 = 2$ kΩ and $R_2 = 18$ kΩ. Sketch and label the magnitude Bode plot of its gain if the circuit also contains a 10-nF capacitance in parallel with R_2.

3.14 Suppose the inverting amplifier of Fig. 1.11 also has a capacitance C_1 in parallel with R_1 and a capacitance C_2 in parallel with R_2. Derive its transfer function, sketch and label

the magnitude Bode plot, and specify suitable component values for a low-frequency gain of 40 dB, a high-frequency gain of 0 dB, and so that the geometric mean of its pole and zero frequencies $(f_p f_z)^{1/2}$ is 1 kHz.

3.15 Sketch and label the linearized magnitude Bode plot for the circuit of Fig. P3.5 if: (*a*) $R_2C_2 = 1$ ms and $R_1C_1 = 0.1$ ms. (*b*) Repeat, but with $R_1C_1 = 10$ ms.

3.16 In the wideband band-pass filter of Fig. 3.11*a* let $R_1 = R_2 = R$ and $C_1 = C_2 = C$. (*a*) Find the output $v_o(t)$ if the input is $v_i(t) = 1\cos(t/RC)$ V. (*b*) Repeat, but for $v_i(t) = 1\cos(t/2RC)$ V. (*c*) Repeat, but for $v_i(t) = 1\cos(t/0.5RC)$ V.

3.17 The circuit of Fig. P3.17 is a capacitance multiplier. (*a*) Show that $C_{eq} = (1+R_2/R_1)C$. (*b*) Using a 0.1-μF capacitance, specify component values to simulate a variable capacitance from 0.1 μF to 100 μF by means of a 1-MΩ pot. *Hint:* In part (*a*), apply a test voltage V, find the resulting current I, and obtain C_{eq} as $1/sC_{eq} = V/I$.

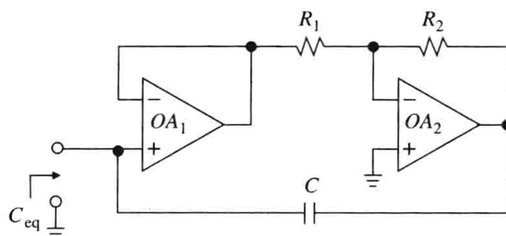


FIGURE P3.17

3.18 The circuit of Fig. P3.18 is a capacitance simulator. (*a*) Show that $C_{eq} = (R_2R_3/R_1R_4)C$. (*b*) Using a 1-nF capacitance, specify component values to simulate a 1-mF capacitance. List a possible application of such a large capacitance. *Hint:* See Problem 3.17.

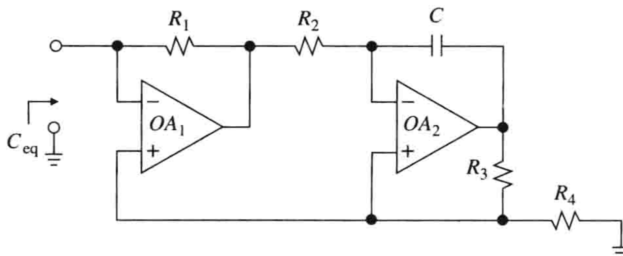


FIGURE P3.18

3.19 In the Deboo integrator of Fig. 3.7 let $R = 16$ kΩ and $2C = 2$ nF. Find the transfer function if the op amp has $a = 10^3$ V/V, $r_d = \infty$, and $r_o = 0$, and verify with PSpice. Compare with the ideal op amp case, and comment.

3.3 Audio filter applications

3.20 Derive Eqs. (3.32) and (3.33).

3.21 (*a*) Derive Eqs. (3.34) and (3.35). (*b*) Specify component values to approximate the NAB curve with a 30-dB gain at 1 kHz. Show the final circuit.

3.22 Using standard component values, design an octave equalizer with center frequencies at approximately $f_0 = 32$ Hz, 64 Hz, 125 Hz, 250 Hz, 500 Hz, 1 kHz, 2 kHz, 4 kHz, 8 kHz, and 16 kHz. Show the final circuit.

3.4 Standard second-order responses

3.23 (*a*) By proper manipulation, put the wideband band-pass function of Eq. (3.29*a*) in the standard form $H(j\omega) = H_{0\text{BP}} H_{\text{BP}}$. (*b*) Show that no matter how you select ω_L and ω_H, the Q of that filter can never exceed $\frac{1}{2}$. This is why the filter is called *wideband.*

3.24 Construct the phase plots of H_{LP}, H_{HP}, H_{BP}, and H_{N} for $Q = 0.2$, 1, and 10.

3.25 (*a*) Use PSpice to plot the transfer function $|V_o/V_i|$ for the circuit of Fig. 3.3*a*, and justify using physical insight. (*b*) Repeat, but for the case in which the output is taken across C. (*c*) Repeat, but for the case in which the output is taken across L. (*d*) Repeat, but for the case in which the output is taken across the L-C combination.

3.26 Suppose $v_i(t)$ in the circuit of Fig. 3.3*a* consists of three ac components, $v_i(t) = 1\sin(0.1\omega_0 t) + 1\sin(\omega_0 t) + 1\sin(10\omega_0 t)$ V, where $\omega_0 = 1/\sqrt{LC}$. (*a*) Use PSpice to plot both $v_i(t)$ and $v_o(t)$, compare, and comment. (*b*) Repeat, but for the case in which $v_o(t)$ is taken across C. (*c*) Repeat, but for the case in which $v_o(t)$ is taken across L. (*d*) Repeat, but for the case in which $v_o(t)$ is taken across the L-C combination.

3.5 *KRC* filters

3.27 An alternative design procedure for the low-pass *KRC* filter of Fig. 3.23 is $R_A = R_B$ and $R_2/R_1 = C_1/C_2 = Q$. (*a*) Develop design equations for this option. (*b*) Hence, use it to redesign the filter of Example 3.8.

3.28 An alternative design procedure for the low-pass *KRC* filter of Fig. 3.23 that allows us to specify also $H_{0\text{LP}}$, $H_{0\text{LP}} > 2$ V/V, is $C_1 = C_2 = C$. (*a*) Show that the design equations for this option are $R_2 = [1 + \sqrt{1 + 4Q^2(H_{0\text{LP}} - 2)}]/2\omega_0 QC$ and $R_1 = 1/\omega_0^2 R_2 C^2$. (*b*) Use this option to redesign the filter of Example 3.8, but with $H_{0\text{LP}} = 10$ V/V.

3.29 (*a*) Design a high-pass *KRC* filter with $f_0 = 100$ Hz and Q variable from 0.5 to 5 by means of a 100-kΩ potentiometer. (*b*) If the input is a 60-Hz, 5-V (rms) ac wave with a dc component of 3 V, what comes out of the filter with the wiper at either extreme?

3.30 An alternative design procedure for the high-pass *KRC* filter of Fig. 3.26 that allows us to specify also $H_{0\text{HP}}$, $H_{0\text{HP}} > 1$, is $C_1 = C_2 = C$. (*a*) Show that the design equations are then $R_1 = [1 + \sqrt{1 + 8Q^2(H_{0\text{HP}} - 1)}]/4\omega_0 QC$ and $R_2 = 1/\omega_0^2 R_1 C^2$. (*b*) Use this option to implement a high-pass Butterworth response with $H_{0\text{HP}} = 10$ V/V and $f_0 = 1$ kHz.

3.31 An alternative design procedure for the band-pass *KRC* filter of Fig. 3.27 is $R_A = R_B$ and $C_1 = C_2 = C$. Develop design equations for this option. Hence, use it to design a band-pass filter with $H_{0\text{BP}} = 0$ dB, $f_0 = 1$ kHz, and $Q = 5$.

3.32 The low-pass filter of Fig. P3.32 is referred to as a $-$*KRC filter* ("minus" *KRC* filter) because the op amp is operated as an inverting amplifier with a gain of $-K$. (*a*) Find $H_{0\text{LP}}$, ω_0, and Q for the case $C_1 = C_2 = C$ and $R_1 = R_2 = R_3 = R_4 = R$. (*b*) Design a $-$*KRC* low-pass filter with $f_0 = 2$ kHz, $Q = 5$, and 0-dB dc gain.

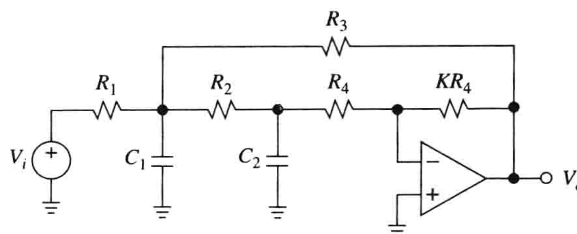


FIGURE P3.32

3.33 The band-pass filter of Fig. P3.33 is referred to as a $-KRC$ *filter* ("minus" KRC filter) because the op amp is operated as an inverting amplifier with a gain of $-K$. (a) Find $H_{0\mathrm{BP}}$, ω_0, and Q for the case $C_1 = C_2 = C$ and $R_1 = R_2 = R$. (b) Design a $-KRC$ band-pass filter with $f_0 = 1$ kHz, $Q = 10$, and unity-resonance gain.

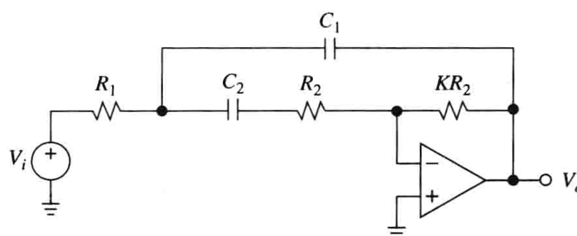


FIGURE P3.33

3.34 The notch filter of Fig. P3.34 allows Q tuning via the ratio R_2/R_1. (a) Show that $V_o/V_i = H_{\mathrm{N}}$ with $\omega_0 = 1/RC$ and $Q = (1 + R_1/R_2)/4$. (b) Specify component values for $f_0 = 60$ Hz and $Q = 25$.

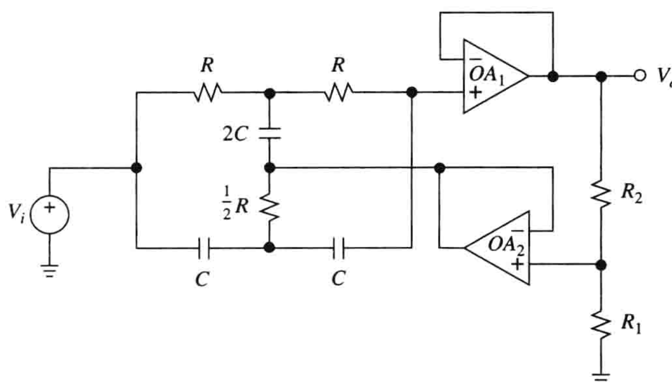


FIGURE P3.34

3.6 Multiple-feedback filters

3.35 An alternative design procedure for the multiple-feedback low-pass filter of Fig. 3.32 is $R_1 = R_2 = R_3 = R$. Find expressions for $H_{0\mathrm{LP}}$, ω_0, and Q. Hence, develop the design equations.

3.36 In the circuit of Fig. 3.33 let $R_3 = R_4 = R$, and $R_5 = KR$. (*a*) Show that if $H_{0BP} = -2$ V/V, the circuit gives the all-pass response with gain $-K$. (*b*) Specify component values for $f_0 = 1$ kHz, $Q = 5$, and a gain of 20 dB.

3.37 Show that the circuit of Fig. P3.37 realizes the all-pass function with $H_{0AP} = 1/3$, $\omega_0 = \sqrt{2}/RC$, and $Q = 1/\sqrt{2}$.

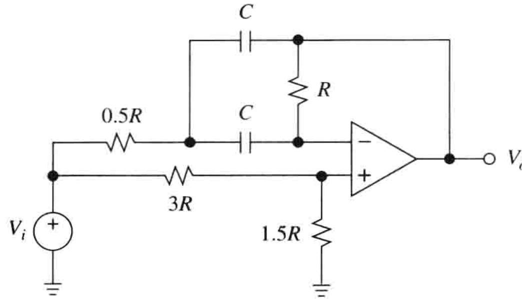


FIGURE P3.37

3.38 The circuit of Fig P3.38, known as a *Q multiplier,* uses a summing amplifier OA_1 and a band-pass stage OA_2 to increase the Q of the band-pass stage without changing ω_0. This allows for high Qs without unduly taxing OA_2. (*a*) Show that the gain and Q of the composite circuit are related to those of the basic band-pass stage as $Q_{comp} = Q/[1-(R_5/R_4)|H_{0BP}|]$, and $H_{0BP(comp)} = (R_5/R_3)(Q_{comp}/Q)H_{0BP}$. (*b*) Specify component values for $f_0 = 3600$ Hz, $Q_{comp} = 60$, and $H_{0BP(comp)} = 2$ V/V, starting with $Q = 10$.

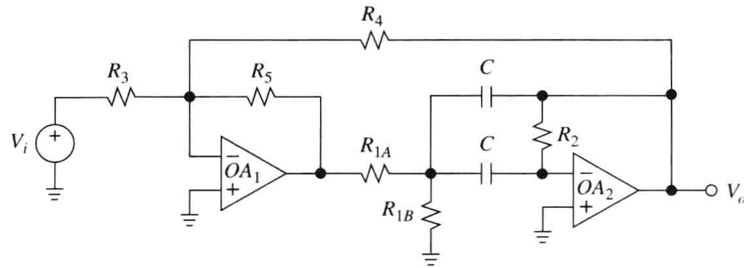


FIGURE P3.38

3.39 With reference to the multiple-feedback low-pass filter of Fig. 3.32, show that the circuit consisting of R_2, R_3, C_2, and the op amp acts as a resistance $R_{eq} = R_2 \parallel R_3$ and an inductance $L_{eq} = R_2R_3C_2$, both in parallel with C_1. Hence, explain circuit operation in terms of the above equivalence.

3.40 Let the circuit of Fig. 3.30 have $C_1 = C_2 = 10$ nF, $R_1 = 10$ kΩ, and $R_2 = 160$ kΩ. Find V_1/V_i, and calculate H_0, ω_0, and Q.

3.41 The multiple-feedback circuit of Fig. P3.41 uses positive feedback to control the Q without upsetting f_0 (see *EDN*, May 11, 1989, p. 200). Show that $Q = 1/(2 - R_1/R_2)$. What are the expressions for f_0 and H_{0BP}? Under what condition is $Q \to \infty$? $Q < 0$? What are the s-plane locations of the poles?

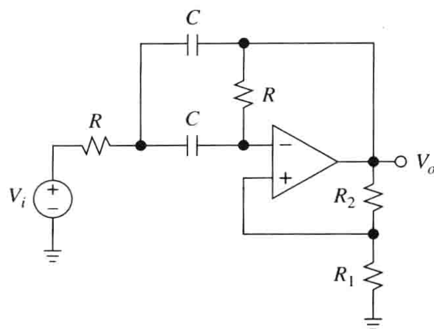


FIGURE P3.41

3.7 State-variable and biquad filters

3.42 Suitably modify the filter of Example 3.18 so that $H_{0\text{BP}} = 1$ V/V. Show your final design.

3.43 (*a*) Derive Eqs. (3.82*a*) and (3.82*b*). (*b*) Specify suitable component values to achieve a band-pass response with $f_L = 594$ Hz and $f_H = 606$ Hz. (*c*) What is the dc gain of the low-pass response?

3.44 The simplified state-variable filter of Fig. P3.44 provides the low-pass and band-pass responses using only two op amps. (*a*) Show that $H_{0\text{BP}} = -n$, $H_{0\text{LP}} = m/(m+1)$, $Q = \sqrt{n(1+1/m)}$, and $\omega_0 = Q/nRC$. (*b*) Specify component values for a band-pass response with $f_0 = 2$ kHz and $Q = 10$. (*c*) What is the resonance gain of your circuit? What is the most serious drawback of this circuit?

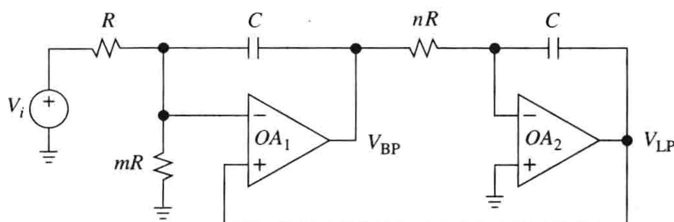


FIGURE P3.44

3.45 Use the noninverting state-variable filter with an additional op amp adder to synthesize the low-pass notch of Example 3.20. *Hint:* Obtain the notch response as $V_{\text{N}} = A_L V_{\text{LP}} + A_H V_{\text{HP}}$, where A_L and A_H are suitable coefficients.

3.46 Consider the dual-op-amp biquad obtained from the standard biquad of Fig. 3.36 by replacing OA_2 and OA_3 with the Deboo integrator of Fig. 3.7. Find its responses; specify component values for a low-pass response with $f_0 = 10$ kHz, $Q = 5$, and $H_{0\text{LP}} = 0$ dB.

3.47 Using the state-variable filter, along with a fourth op amp adder, design an all-pass circuit with $f_0 = 1$ kHz and $Q = 1$. *Hint:* Apply Eq. (3.58).

3.48 Consider the biquad of Fig. 3.36 with R_2 removed. If the op amps are ideal, the circuit will give $Q \to \infty$, but what if the op amps have finite gains? Assuming equal capacitors

and equal resistors throughout, find Q if the op amps have $a = 100$ V/V, $r_d = \infty$, and $r_o = 0$. Verify with PSpice.

3.8 Sensitivity

3.49 Prove Eq. (3.92).

3.50 Show that any second-order *KRC* filter, in which K appears only in the s-term in the denominator, has always $S_K^Q > 2Q - 1$.

3.51 An alternative design procedure for the multiple-feedback low-pass filter of Fig. 3.32 is $R_1 = R_2 = R_3 = R$. (*a*) Find simplified expressions for ω_0 and Q. (*b*) Find the sensitivity functions.

3.52 Calculate the sensitivities of the state-variable filter of Example 3.18.

3.53 Let the circuit of Fig. P3.41 be implemented with $R = 1.59$ kΩ, $C = 1$ nF, and $R_1 = R_2 = 10$ kΩ. Use PSpice to find how a 10% change in each component, one component at a time, affects the resonant frequency f_0. Is this a good circuit in terms of sensitivities?

REFERENCES

1. M. E. Van Valkenburg, *Analog Filter Design,* Holt, Rinehart and Winston, Orlando, FL, 1982.
2. L. P. Huelsman, *Active and Passive Analog Filter Design: An Introduction,* McGraw-Hill, New York, 1993.
3. F. W. Stephenson, *RC Active Filter Design Handbook,* John Wiley & Sons, New York, 1985.
4. A. B. Williams and F. J. Taylor, *Electronic Filter Design Handbook: LC, Active, and Digital Filters,* 2d ed., McGraw-Hill, New York, 1988.
5. S. Franco, *Electric Circuits Fundamentals,* Oxford University Press, New York, 1995.
6. W. G. Jung, *Audio IC Op Amp Applications,* 3d ed., Howard W. Sams, Carmel, IN, 1987.
7. K. Lacanette, "High Performance Audio Applications of the LM833," Application Note AN-346, *Linear Applications Handbook,* National Semiconductor, Santa Clara, CA, 1994.
8. R. A. Greiner and M. Schoessow, "Design Aspects of Graphic Equalizers," *J. Audio Eng. Soc.,* Vol. 31, No. 6, June 1983, pp. 394–407.

4

ACTIVE FILTERS: PART II

前面我们已经学习了一阶和二阶滤波器，现在开始研究高阶滤波器。当低阶滤波器的截止特性的陡峭程度不能够满足实际应用需求的时候，就需要高阶滤波器。在众多高阶滤波器的实现方法中，级联设计和直接综合设计方法应用最为广泛。级联设计方法通过在第 3 章中学习过的二阶滤波器（也可能是一阶滤波器）级联起来实现希望获得的响应。直接综合设计方法利用诸如回旋器和频率负阻（frequency-dependent negative resistance）等阻抗变换器来模拟原型满足给定指标特性要求的无源 *RLC* 滤波器。

不考虑它们响应的复杂程度，以上滤波器都称为连续时间滤波器。因为电路本身采用大体积的电容以及对于 RC 元件所控制的频率特性的精度和稳定度具有严格要求，所以这些滤波器都无法制造成单片结构。另一方面，现代的超大规模集成电路（VLSI）通常将数字功能和模拟功能集成在同一片芯片上。为了满足滤波或者其他传统的模拟电路功能，开关电容技术已经得到了长足发展。它采用 MOS 运算放大器、电容和开关，而不使用电阻，在一个相对有限的频率范围内实

现相当稳定的滤波功能。开关电容（SC）电路属于数据采样系统。这种系统所处理的信息是在离散的时间点上进行的，而不是在连续的。这将限制它在音频领域的应用。

本章重点

本章第一部分首先回顾了通常使用的滤波器近似。之后讨论了两种最为广泛使用的滤波器设计方法——级联设计和直接综合设计。第二部分首先介绍了开关电容技术，之后研究了开关电容滤波器，并且说明了它的应用和限制。如同在第 3 章，为了只将注意力集中在滤波器的概念而不考虑因运算放大器的非理想特性引起的复杂性（运算放大器的动态限制效应将在第 6 章讨论），我们故意假设运算放大器是理想的。像第 3 章一样，SPICE 为滤波器响应的可视化和像示波器一样的测量滤波器特性提供了有力的工具。

4.1 FILTER APPROXIMATIONS

If the signals to be rejected are very close in frequency to those that must be passed, the cutoff characteristics of a second-order filter may not prove sufficiently sharp, so a higher-order filter may be needed. Actual filters can only approximate the brick-wall responses of Fig. 3.1. In general, the closer the desired approximation, the higher the order of the filter.

The departure of a practical filter from its brick-wall model is visualized in terms of a shaded area,[1] as shown in Fig. 4.1a for the low-pass case. Introducing the attenuation $A(\omega)$ as

$$A(\omega) = -20\log_{10}|H(j\omega)| \tag{4.1}$$

we observe that the range of frequencies that are passed with little or no attenuation defines the *passband*. For a low-pass filter, this band extends from dc to some frequency ω_c, called the *cutoff frequency*. Gain is not necessarily constant within the

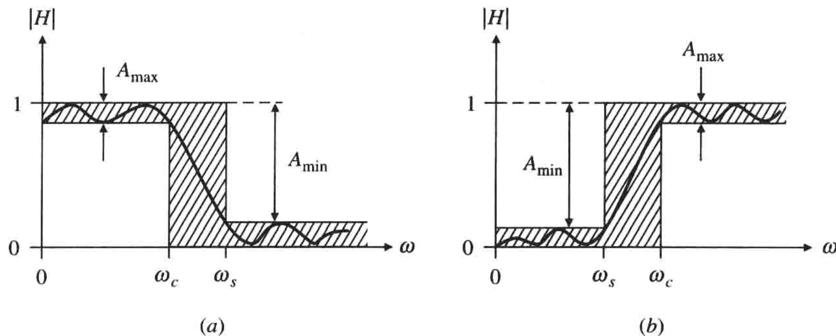


FIGURE 4.1
Magnitude limits for (a) the low-pass and (b) the high-pass responses.

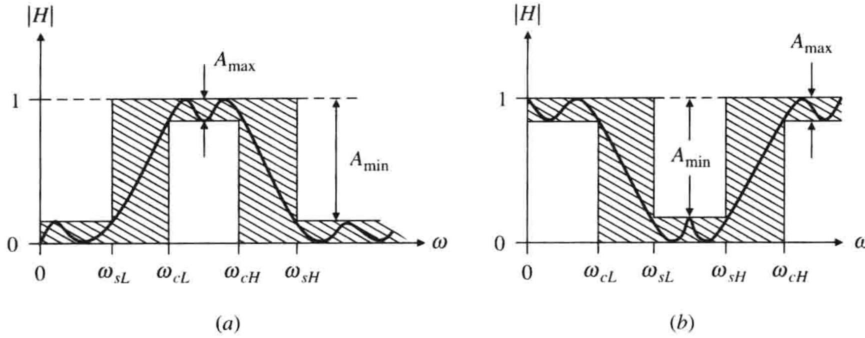


FIGURE 4.2
Magnitude limits for (*a*) the band-pass and (*b*) the band-reject responses.

passband but is allowed a maximum variation $A_{\max}$, such as $A_{\max} = 1$ dB. Gain may exhibit ripple within the passband, in which case $A_{\max}$ is called the *maximum passband ripple* and the passband is called the *ripple band.* In this case ω_c represents the frequency at which the response departs from the ripple band.

Past ω_c the magnitude drops off to the *stopband,* or the frequency region of substantial attenuation. This band is specified in terms of some minimum allowable attenuation, such as $A_{\min} = 60$ dB. The frequency at which the stopband begins is denoted as ω_s. The ratio ω_s/ω_c is called the *selectivity factor* because it gives a measure of the sharpness of the response. The frequency region between ω_c and ω_s is called the *transition band,* or *skirt.* Certain filter approximations maximize the rate of descent within this band at the expense of ripples within the other bands.

The terminology developed for the low-pass case is readily extended to the high-pass case depicted in Fig. 4.1*b*, and to the band-pass and band-reject cases depicted in Fig. 4.2.

As the order n of a transfer function is increased, additional parameters are brought into play in the form of the higher-order polynomial coefficients. These coefficients provide the designer with additional freedom in specifying the frequency profiles of magnitude or phase, thus allowing for an increased degree of optimization. Among the various approximations, some have been found to be consistently satisfactory to justify the tabulation of their coefficients in filter handbooks. These include the *Butterworth, Chebyshev, Cauer,* and *Bessel* approximations.

Filter tables list the denominator polynomial coefficients of the various approximations for a cutoff frequency of 1 rad/s. As an example, the coefficients of the fifth-order Butterworth response are[2] $b_0 = b_5 = 1$, $b_2 = b_4 = 3.236$, and $b_3 = 5.236$, so

$$H(s) = \frac{1}{s^5 + 3.236s^4 + 5.236s^3 + 5.236s^2 + 3.236s + 1} \tag{4.2}$$

An alternative approach is to factor out $H(s)$ into the product of terms of order ≤ 2 and tabulate the coefficients of these terms instead. Expressed in this form, the above function becomes

$$H(s) = \frac{1}{s^2 + 0.6180s + 1} \times \frac{1}{s^2 + 1.6180s + 1} \times \frac{1}{s + 1} \tag{4.3}$$

The design of a higher-order filter begins with the selection of the approximation best suited to the given application, followed by the specification of ω_c, ω_s, $A_{\max}$, and $A_{\min}$. The latter are then used as keys to filter handbooks or computer programs to find the required order n. Once n is known, various alternatives are available to the active-filter designer, the most popular ones being the *cascade* approach and the *RLC ladder simulation* approach. The cascade approach realizes the desired response by cascading lower-order stages of the type investigated in Chapter 3. The ladder simulation approach utilizes active impedance converters, such as gyrators and frequency-dependent negative resistors, to simulate a passive *RLC* filter prototype meeting the desired specifications.

Once an approach has been chosen, one must find the individual-stage values of ω_0 and Q (and possibly ω_z) in the case of cascade design, or the individual values of R, L, and C in the case of ladder simulation. These data are again found with the help of filter tables or computer programs. To promote their products, analog companies offer a variety of filter-design tools that the user can download for free from the web (in this respect, you are encouraged to perform a web search for "filter design software" and secure your own filter program). In the following discussions we shall continue to use FILDES, one of the earliest cascade-design programs developed for personal computers, which has proven helpful in previous editions of this book, and which the reader can download from this book's website, http://www.mhhe.com/franco.

Plotting $H(jf)$ Using PSpice

As already seen in Chapter 3, a convenient means for generating the Bode plots of a filter function $H(s)$ is PSpice's Laplace block. As an example, Fig. 4.3 shows a PSpice circuit to plot the magnitude of the fifth-order Butterworth response of Eq. (4.3). Magnitude is plotted in Fig. 4.4 using linear as well as logarithmic scales, so we can better appreciate similarities and differences (the linear plot is preferable in relation to Fig. 4.1, whereas the logarithmic plot affords a much better visualization of high- and low-frequency details that get far too compressed in the linear plot). Note, in the logarithmic plot, the rapid descent from 0 dB to –100 dB within a single decade, as expected of a fifth-order filter.

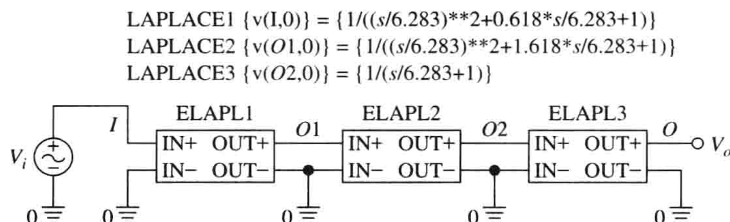


FIGURE 4.3
PSpice circuit to plot the fifth-order Butterworth low-pass function normalized to 1 Hz.

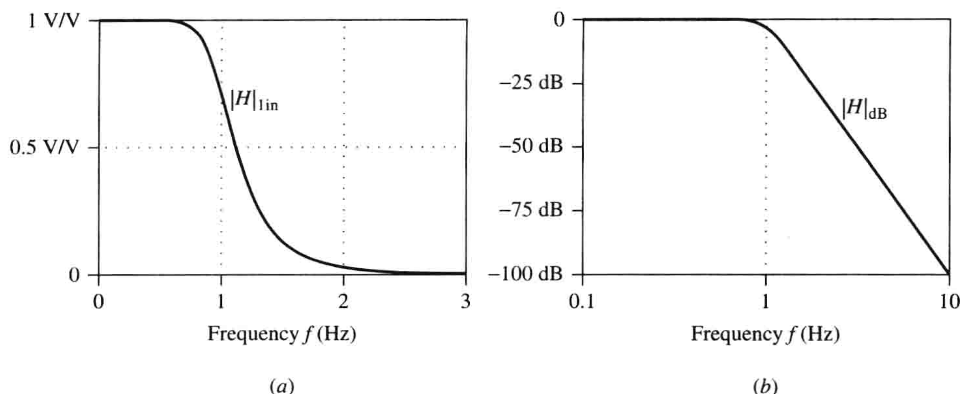


FIGURE 4.4
Magnitude plots of the fifth-order Butterworth low-pass function normalized to 1 Hz. (*a*) Linear scale and (*b*) logarithmic scale.

Butterworth Approximation

The gain of the Butterworth approximation is[3]

$$|H(j\omega)| = \frac{1}{\sqrt{1+\epsilon^2(\omega/\omega_c)^{2n}}} \tag{4.4}$$

where n is the order of the filter, ω_c is the cutoff frequency, and ϵ is a constant that determines the maximum passband variation as $A_{\max} = A(\omega_c) = 20 \times \log_{10}\sqrt{1+\epsilon^2} = 10\log_{10}(1+\epsilon^2)$. The first $2n-1$ derivatives of $|H(j\omega)|$ are zero at $\omega = 0$, indicating a curve as flat as possible at $\omega = 0$. Aptly referred to as *maximally flat,* a Butterworth curve becomes somewhat rounded near ω_c and rolls off at an ultimate rate of $-20n$ dB/dec in the stopband. As shown in Fig. 4.5*a* for $\epsilon = 1$, the higher the order n, then the closer the response is to the brick-wall model.

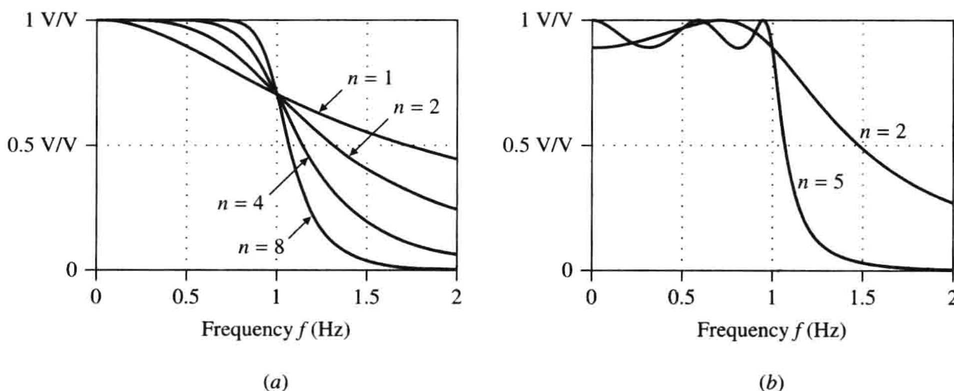


FIGURE 4.5
(*a*) Butterworth and (*b*) 1-dB Chebyshev responses.

EXAMPLE 4.1. Find n for a low-pass Butterworth response with $f_c = 1$ kHz, $f_s = 2$ kHz, $A_{max} = 1$ dB, and $A_{min} = 40$ dB.

Solution. Letting $A_{max} = A(\omega_c) = 20\log_{10}\sqrt{1+\epsilon^2} = 1$ dB gives $\epsilon = 0.5088$. Letting $A(\omega_s) = 10\log_{10}[1+\epsilon^2(2/1)^{2n}] = 40$ dB, we find that $n = 7$ gives $A(\omega_s) = 36.3$ dB and $n = 8$ gives $A(\omega_s) = 42.2$ dB. For $A_{min} = 40$ dB we thus select $n = 8$.

Chebyshev Approximation

There are applications where sharp cutoff is more important than maximal flatness. Chebyshev filters maximize the transition-band cutoff rate at the price of introducing passband ripples, as shown in Fig. 4.5*b*. As a general rule, the higher A_{max}, the narrower the transition band for a given A_{min}. The gain of an nth-order Chebyshev approximation with cutoff frequency ω_c and $A_{max} = 10\log_{10}(1+\epsilon^2)$ is[3]

$$|H(j\omega)| = \frac{1}{\sqrt{1+\epsilon^2 C_n^2(\omega/\omega_c)}} \tag{4.5}$$

where $C_n(\omega/\omega_c)$ is the Chebyshev polynomial of order n, defined as

$$C_n(\omega/\omega_c \le 1) = \cos[n\cos^{-1}(\omega/\omega_c)] \tag{4.6a}$$

$$C_n(\omega/\omega_c \ge 1) = \cosh[n\cosh^{-1}(\omega/\omega_c)] \tag{4.6b}$$

We observe that $C_n^2(\omega/\omega_c \le 1) \le 1$, and $C_n^2(\omega/\omega_c \ge 1) \ge 1$. Moreover, within the passband $|H(j\omega)|$ exhibits peak values of 1 and valley values of $1/\sqrt{1+\epsilon^2}$ at the frequencies that make the cosine term zero and unity, respectively. The number of these peaks or valleys, including the one at the origin, is n.

Compared to the Butterworth approximation, which exhibits appreciable departure from its dc value only at the upper end of the passband, the Chebyshev approximation improves the transition-band characteristic by spreading its equal-sized ripples throughout the passband. At dc, the decibel value of a Chebyshev response is 0 if n is odd, and $0 - A_{max}$ if n is even. A Chebyshev filter can achieve a given transition-band cutoff rate with a lower order than a Butterworth filter, thus reducing circuit complexity and cost. Past the transition band, however, the Chebyshev response rolls off at an ultimate rate of $-20n$ dB/dec, just like a Butterworth response of the same order.

Cauer Approximation

Cauer filters, also called *elliptic filters,* carry the Chebyshev approach one step further by trading ripples in both the passband and the stopband for an even sharper characteristic in the transition band. Consequently, they can provide a given transition-band cutoff rate with an even lower order n than Chebyshev filters. The idea is to follow an existing low-pass response with a notch just above ω_c to further sharpen the response. To be effective, the notch must be narrow, indicating that the curve will come back up just past this notch. At this point another notch is created to press the curve back down, and the process is repeated until the overall profile within the stopband is pushed below the level specified by A_{min}. The various approximations are compared in Fig. 4.6 for $n = 5$ and $A_{max} = 3$ dB.

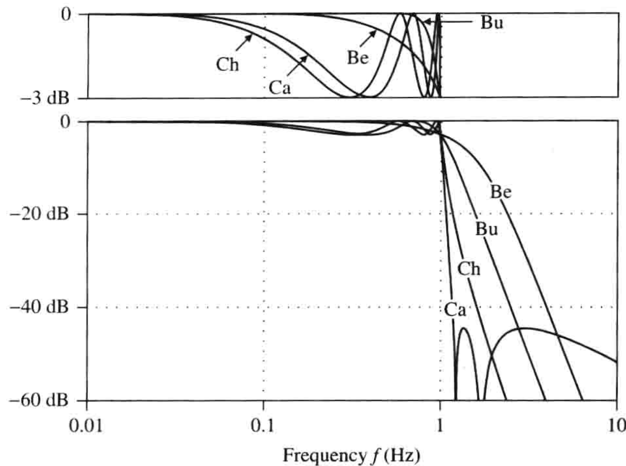


FIGURE 4.6
Fifth-order filter responses comparison: Bessel (Be), Butterworth (Bu), Chebyshev (Ch) and Cauer (Ca). Shown at the top is an expanded view of the 3-dB passband.

Bessel Approximation

In general, filters introduce a frequency-dependent phase shift. If this shift varies linearly with frequency, its effect is simply to delay the signal by a constant amount. However, if phase varies nonlinearly, different input frequency components will experience different delays, so nonsinusoidal signals may experience significant phase distortion in propagating through the filter. In general, the steeper the transition-band magnitude characteristic, the higher the distortion.[1]

Bessel filters, also called *Thomson filters,* maximize the passband delay just as Butterworth filters maximize the passband magnitude. The result is a nearly linear phase characteristic within the passband, if at the price of a less sharp magnitude characteristic in the transition band. Figure 4.7 shows that a pulse emerges fairly undistorted from a Bessel filter, but exhibits appreciable overshoot and ringing when processed with a Chebyshev filter, whose phase response is less linear than Bessel's.

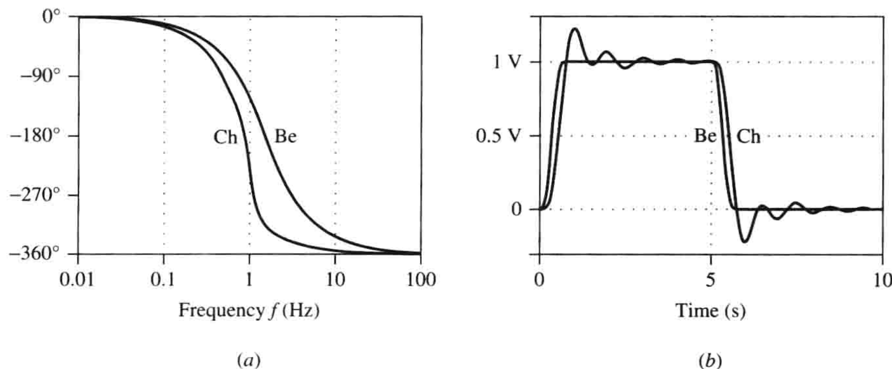


FIGURE 4.7
Comparing the (*a*) phase and (*b*) pulse responses of fourth-order Bessel (Be) and 1-dB Chebyshev (Ch) filters.

4.2 CASCADE DESIGN

This approach is based on the factorization of a transfer function $H(s)$ into the product of lower-order terms. If the order n is even, the decomposition consists of $n/2$ second-order terms,

$$H(s) = H_1(s) \times H_2(s) \times \cdots \times H_{n/2}(s) \tag{4.7}$$

If n is odd, the factorization includes also a first-order term. Sometimes this term is combined with one of the second-order terms to create a third-order filter stage. The first-order term, if any, can be implemented with a plain RC or CR network, so all we need to know is its required frequency ω_0. The second-order terms can be implemented with any of the filters of Sections 3.5 through 3.7. For each of these stages we need to know ω_0 and Q, and ω_z if the stage is a notch stage. As mentioned, these data are tabulated in filter handbooks[3] or can be calculated by computer.[4]

The cascade approach offers a number of advantages. The design of each section is relatively simple, and the component count is usually low. The low-output impedance of the individual sections eliminates interstage loading, so each section can be regarded as isolated from the others and can be tuned independently, if needed. The inherent modularity of this approach is also attractive from the economic standpoint, since one can use a few standardized blocks to design a variety of more complex filters.

Mathematically, the order in which the various sections are cascaded is irrelevant. In practice, to avoid loss of dynamic range and filter accuracy due to possible signal clipping in the high-Q sections, the sections are cascaded in order of ascending Qs, with the low-Q stages first in the signal path. This ordering, however, does not take into account internal noise, which may be of concern in the high-Q stages, where any noise component falling under the resonance peak may be amplified significantly. So, to minimize noise, high-Q stages should go first in the cascade. In general, the optimum ordering depends on the input spectrum, the filter type, and the noise characteristics of its components.[5]

Low-Pass Filter Design

Table 4.1 gives examples of tabulated data for cascade design. Butterworth and Bessel data are tabulated for different values of n, Chebyshev data for different values of n and $A_{\max}$ (shown in the table are the data for $A_{\max} = 0.1$ dB and $A_{\max} = 1.0$ dB), and Cauer data (not shown in the table) for different values of n, $A_{\max}$, and $A_{\min}$. Frequency data are expressed in normalized form for a cutoff frequency of 1 Hz. In the Butterworth and Bessel cases this frequency coincides with the -3-dB frequency, while in the Chebyshev and Cauer cases it represents the frequency at which gain departs from the ripple band. To convert from normalized to actual frequencies, we simply multiply the tabulated values by the cutoff frequency f_c of the filter being designed, or

$$f_0 = f_{0(\text{table})} \times f_c \tag{4.8a}$$

In the case of Cauer filters, the tables include not only pole frequencies but also zero frequencies. The latter are converted as

$$f_z = f_{z(\text{table})} \times f_c \tag{4.8b}$$

TABLE 4.1
Examples of normalized (1 Hz) low-pass filter tables

Butterworth low-pass filter

n	f_{01}	Q_1	f_{02}	Q_2	f_{03}	Q_3	f_{04}	Q_4	f_{05}	Q_5	Att (dB) at $2f_c$
2	1	0.707	1								12.30
3	1	1.000	1								18.13
4	1	0.541	1	1.306							24.10
5	1	0.618	1	1.620	1						30.11
6	1	0.518	1	0.707	1	1.932					36.12
7	1	0.555	1	0.802	1	2.247	1				42.14
8	1	0.510	1	0.601	1	0.900	1	2.563			48.16
9	1	0.532	1	0.653	1	1.000	1	2.879	1		54.19
10	1	0.506	1	0.561	1	0.707	1	1.101	1	3.196	60.21

Bessel low-pass filter

n	f_{01}	Q_1	f_{02}	Q_2	f_{03}	Q_3	f_{04}	Q_4	f_{05}	Q_5
2	1.274	0.577								
3	1.453	0.691	1.327							
4	1.419	0.522	1.591	0.806						
5	1.561	0.564	1.760	0.917	1.507					
6	1.606	0.510	1.691	0.611	1.907	1.023				
7	1.719	0.533	1.824	0.661	2.051	1.127	1.685			
8	1.784	0.506	1.838	0.560	1.958	0.711	2.196	1.226		
9	1.880	0.520	1.949	0.589	2.081	0.760	2.324	1.322	1.858	
10	1.949	0.504	1.987	0.538	2.068	0.620	2.211	0.810	2.485	1.415

0.10-dB ripple Chebyshev low-pass filter

n	f_{01}	Q_1	f_{02}	Q_2	f_{03}	Q_3	f_{04}	Q_4	f_{05}	Q_5	Att (dB) at $2f_c$
2	1.820	0.767									3.31
3	1.300	1.341	0.969								12.24
4	1.153	2.183	0.789	0.619							23.43
5	1.093	3.282	0.797	0.915	0.539						34.85
6	1.063	4.633	0.834	1.332	0.513	0.599					46.29
7	1.045	6.233	0.868	1.847	0.575	0.846	0.377				57.72
8	1.034	8.082	0.894	2.453	0.645	1.183	0.382	0.593			69.16
9	1.027	10.178	0.913	3.145	0.705	1.585	0.449	0.822	0.290		80.60
10	1.022	12.522	0.928	3.921	0.754	2.044	0.524	1.127	0.304	0.590	92.04

1.00-dB ripple Chebyshev low-pass filter

n	f_{01}	Q_1	f_{02}	Q_2	f_{03}	Q_3	f_{04}	Q_4	f_{05}	Q_5	Att (dB) at $2f_c$
2	1.050	0.957									11.36
3	0.997	2.018	0.494								22.46
4	0.993	3.559	0.529	0.785							33.87
5	0.994	5.556	0.655	1.399	0.289						45.31
6	0.995	8.004	0.747	2.198	0.353	0.761					56.74
7	0.996	10.899	0.808	3.156	0.480	1.297	0.205				68.18
8	0.997	14.240	0.851	4.266	0.584	1.956	0.265	0.753			79.62
9	0.998	18.029	0.881	5.527	0.662	2.713	0.377	1.260	0.159		91.06
10	0.998	22.263	0.902	6.937	0.721	3.561	0.476	1.864	0.212	0.749	102.50

A common application of low-pass filters is in connection with analog-to-digital (A-D) and digital-to-analog (D-A) conversion. By the well-known sampling theorem, the input signal to an A-D converter must be band limited to less than half the sampling frequency in order to avoid *aliasing*. Likewise, the output signal of a D-A converter must be properly smoothed in order to avoid the effects of discrete quantization and time sampling. Both tasks are accomplished with sharp low-pass filters designed to provide adequate attenuation at half the sampling frequency.

EXAMPLE 4.2. The output of a D-A converter with a sampling rate of 40 kHz is to be smoothed with a sixth-order 1.0-dB Chebyshev low-pass filter providing an attenuation of 40 dB at half the sampling frequency, or 20 kHz. This attenuation requirement is met by letting $f_c = 13.0$ kHz. (*a*) Design such a filter. (*b*) Verify with PSpice.

Solution.

(*a*) From Table 4.1 we find that a 1.0-dB Chebyshev filter with $n = 6$ requires three second-order stages with

$$f_{01} = 0.995 f_c = 12.9 \text{ kHz} \qquad Q_1 = 8.00$$
$$f_{02} = 0.747 f_c = 9.71 \text{ kHz} \qquad Q_2 = 2.20$$
$$f_{03} = 0.353 f_c = 4.59 \text{ kHz} \qquad Q_3 = 0.761$$

Use three unity-gain Sallen-Key sections and cascade them in order of ascending Qs. Retracing the design steps of Example 3.10, we find the component values shown in Fig. 4.8(top), where the resistances have been rounded off to the nearest 1% standard values.

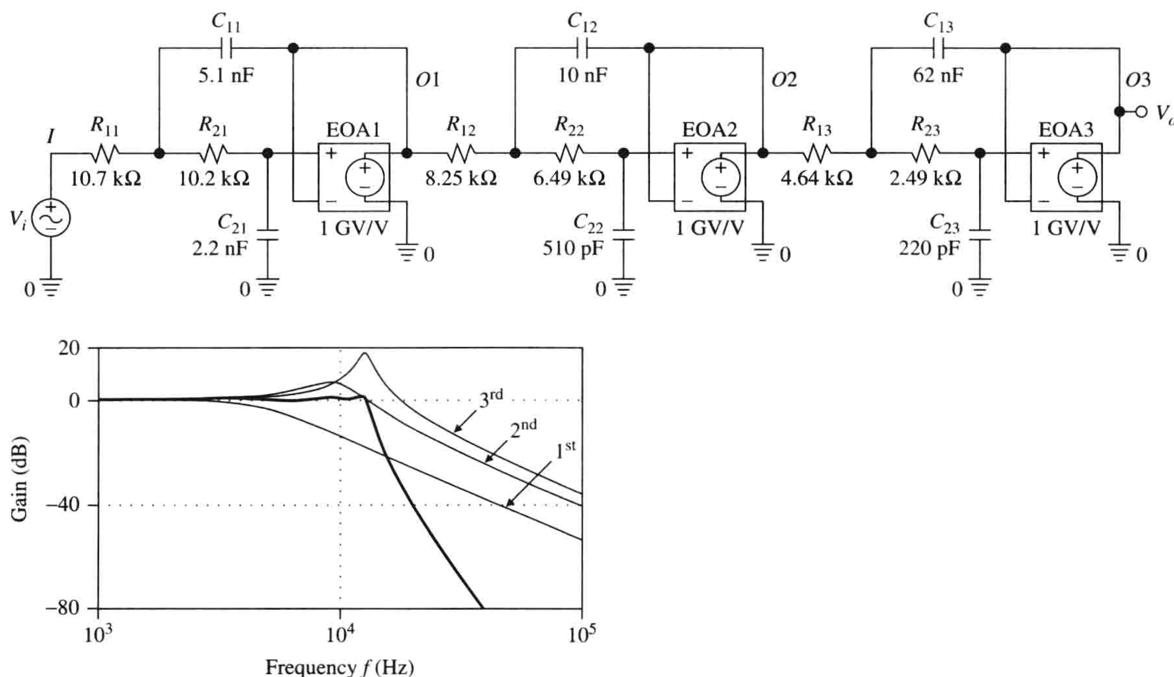


FIGURE 4.8
Top: Sixth-order Chebyshev low-pass filter of Example 4.2. Bottom: Magnitude plot of its overall response $H = V_o/V_i$, along with the responses of its individual stages.

Figure 4.8 (bottom) shows the overall response as well as the individual-stage responses. It is interesting to observe how the latter combine to create the ripple and cutoff characteristics of the overall response.

EXAMPLE 4.3. Design a Cauer low-pass filter with $f_c = 1$ kHz, $f_s = 1.3$ kHz, $A_{max} = 0.1$ dB, $A_{min} = 40$ dB, and dc gain $H_0 = 0$ dB.

Solution. Using the aforementioned filter design program FILDES (check our website for information on how to download this program), we find that a sixth-order implementation is required, with the following individual-stage parameters:

$$f_{01} = 648.8 \text{ Hz} \qquad f_{z1} = 4130.2 \text{ Hz} \qquad Q_1 = 0.625$$

$$f_{02} = 916.5 \text{ Hz} \qquad f_{z2} = 1664.3 \text{ Hz} \qquad Q_2 = 1.789$$

$$f_{03} = 1041.3 \text{ Hz} \qquad f_{z3} = 1329.0 \text{ Hz} \qquad Q_3 = 7.880$$

Moreover, the program indicates that the actual attenuation at 1.3 kHz is 47 dB, and the −3-dB frequency is 1.055 kHz.

We shall implement the filter with three low-pass notch sections of the biquad type of Fig. 3.37. Using Eq. (3.89) and retracing the steps of Example 3.20, we find the component values shown in Fig. 4.9, where the resistances have been rounded off to the nearest 1% standard values. The entire filter can be built with three quad-op-amp packages.

High-Pass Filter Design

Owing to the fact that a high-pass transfer function can be obtained from a low-pass function via the substitution $s/\omega_0 \rightarrow 1/(s/\omega_0)$, the normalized frequency data of Table 4.1 can also be used in the cascade design of high-pass filters, provided actual frequencies are obtained from tabulated frequencies as

$$f_0 = f_c/f_{0(\text{table})} \tag{4.9a}$$

$$f_z = f_c/f_{z(\text{table})} \tag{4.9b}$$

where f_c is the cutoff frequency of the filter being designed.

EXAMPLE 4.4. Design a third-order, 0.1-dB Chebyshev high-pass filter with $f_c = 100$ Hz and high-frequency gain $H_0 = 20$ dB.

Solution. Table 4.1 indicates that we need a second-order high-pass section with $f_{01} = 100/1.300 = 76.92$ Hz and $Q_1 = 1.341$, and a first-order high-pass section with $f_{02} = 100/0.969 = 103.2$ Hz. As shown in Fig. 4.10, we implement the filter with a second-order unity-gain Sallen-Key high-pass stage, followed by a first-order high-pass stage with a high-frequency gain of 10 V/V.

Band-Pass Filter Design

EXAMPLE 4.5. Design a Butterworth band-pass filter with center frequency $f_0 = 1$ kHz, BW = 100 Hz, $A(f_0/2) = A(2f_0) \geq 60$ dB, and resonance gain $H_0 = 0$ dB.

Solution. Using the aforementioned FILDES program, we find that the given specifications can be met with a sixth-order filter having the following individual-stage

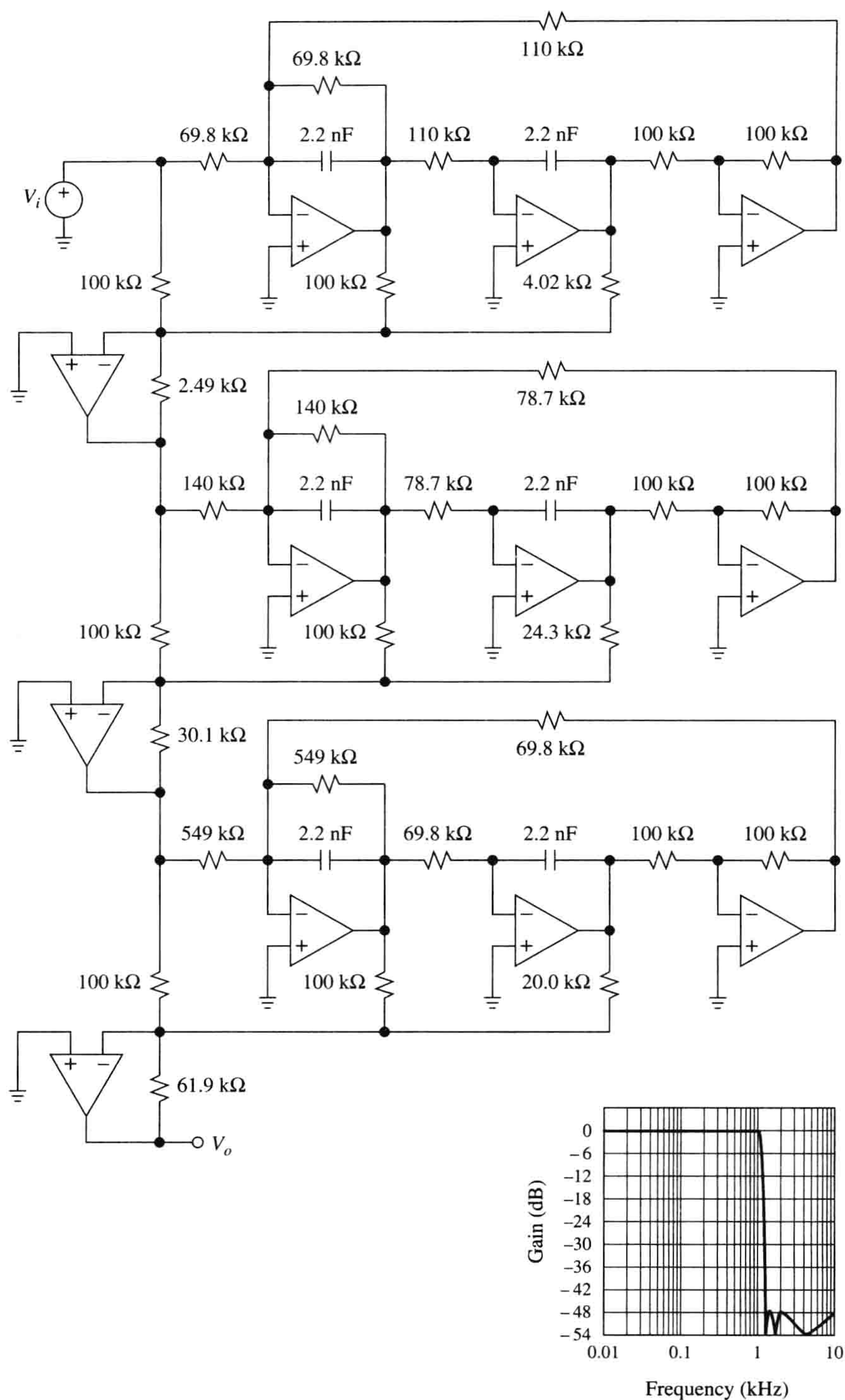


FIGURE 4.9
Sixth-order 0.1/40-dB elliptic low-pass filter of Example 4.3.

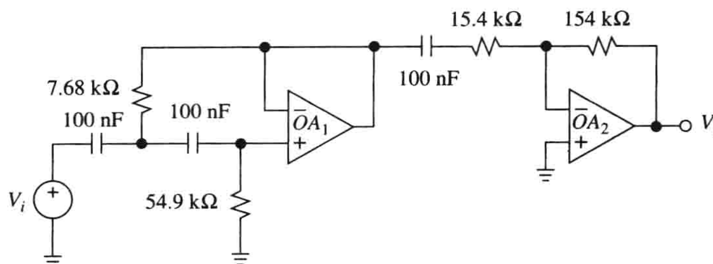


FIGURE 4.10
Third-order 0.1-dB Chebyshev high-pass filter of Example 4.4.

parameters:

$$f_{01} = 957.6 \text{ Hz} \qquad Q_1 = 20.02$$
$$f_{02} = 1044.3 \text{ Hz} \qquad Q_2 = 20.02$$
$$f_{03} = 1000.0 \text{ Hz} \qquad Q_3 = 10.0$$

Furthermore, the actual attenuation at 500 Hz and 2 kHz is 70.5 dB, and the midband gain is -12 dB, that is, 0.25 V/V. To raise it to 0 dB, we shall impose $H_{0\text{BP}1} = H_{0\text{BP}2} = 2$ V/V, and $H_{0\text{BP}3} = 1$ V/V.

We shall implement the filter with three multiple-feedback band-pass sections equipped with input resistance attenuators. Retracing the steps of Example 3.15 we find the components of Fig. 4.11, where the resistances have been rounded off to 1% standard values, and the second leg of each attenuator has been made variable for tuning purposes. To tune a given section, apply an ac input at the desired resonance frequency of that section, and adjust its pot until the Lissajous figure changes from an ellipse to a straight segment.

EXAMPLE 4.6. Design an elliptic band-pass filter with $f_0 = 1$ kHz, passband = 200 Hz, stopband = 500 Hz, $A_{\text{max}} = 1$ dB, $A_{\text{min}} = 40$ dB, and $H_0 = 20$ dB.

Solution. The above-mentioned FILDES program indicates that we need a sixth-order filter with the following individual-stage parameters:

$$f_{01} = 907.14 \text{ Hz} \qquad f_{z1} = 754.36 \text{ Hz} \qquad Q_1 = 21.97$$
$$f_{02} = 1102.36 \text{ Hz} \qquad f_{z2} = 1325.6 \text{ Hz} \qquad Q_2 = 21.97$$
$$f_{03} = 1000.0 \text{ Hz} \qquad Q_3 = 9.587$$

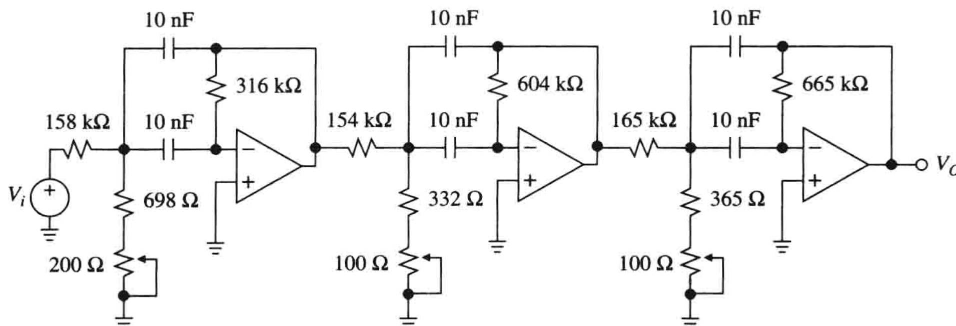


FIGURE 4.11
Sixth-order Butterworth band-pass filter of Example 4.5.

Moreover, the actual attenuation at the stopband edges is 41 dB, and the midband gain is 18.2 dB. We shall implement the filter with a high-pass notch biquad stage, a low-pass notch biquad stage, and a multiple-feedback band-pass stage. To bolster the midband gain from 18.2 dB to 20 dB, we impose $H_{0BP3} = 1.23$ V/V, and to simplify inventory we use 10-nF capacitances throughout.

Using Eq. (3.89) we find, for the high-pass notch, $R = 1/(2\pi \times 907.14 \times 10^{-8}) = 17.54\,\text{k}\Omega$, $R_1 = 21.97 \times 17.54 = 385.4\,\text{k}\Omega$, $R_2 = R_3 = 100\,\text{k}\Omega$, $R_4 = (100/21.97)907.14^2/(907.14^2 - 754.36^2) = 14.755\,\text{k}\Omega$, and $R_5 = 100\,\text{k}\Omega$. Proceeding in like manner for the other two sections, we end up with the circuit of Fig. 4.12, where the resistances have been rounded off to 1% standard values, and provisions have been made for frequency and Q tuning.

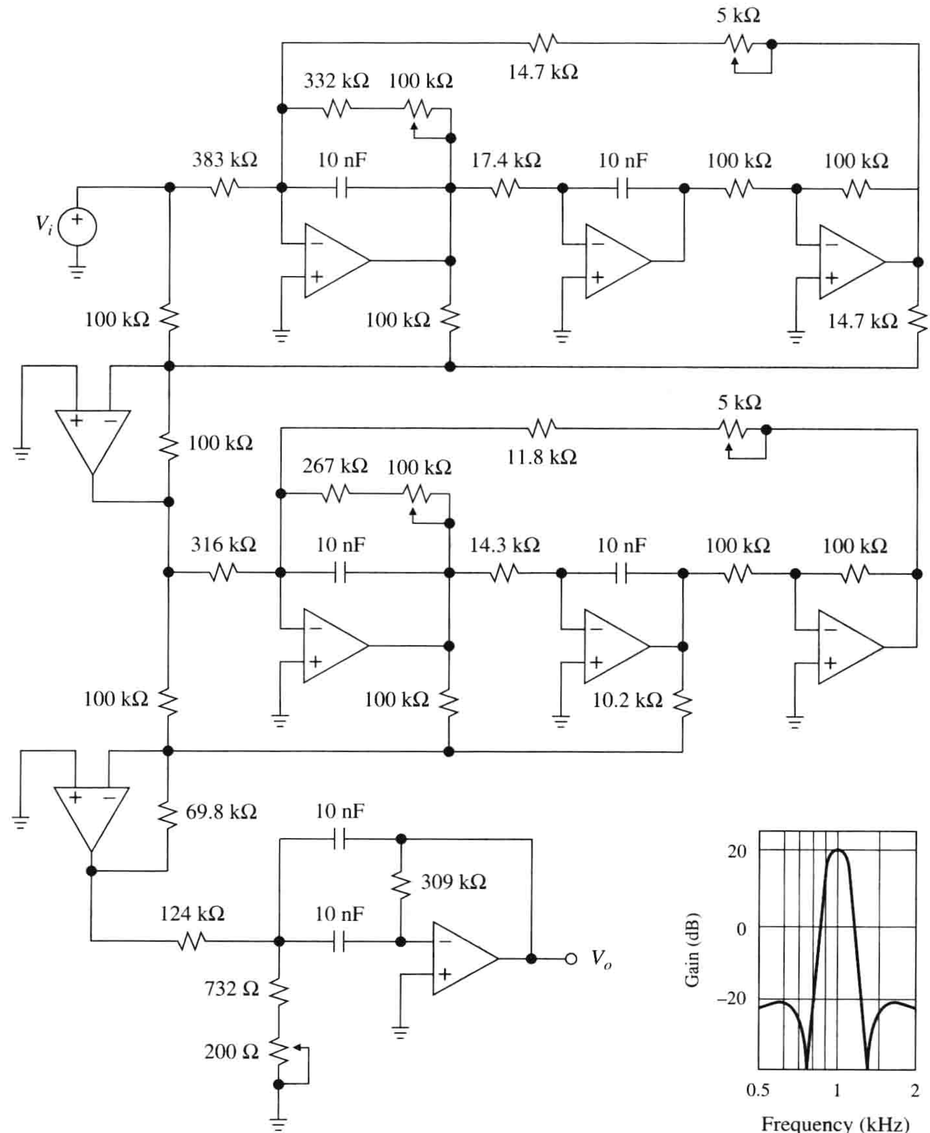


FIGURE 4.12
Sixth-order 1.0/40-dB elliptic band-pass filter of Example 4.6.

EXAMPLE 4.7. A 0.1-dB Chebyshev band-reject filter is to be designed with notch frequency $f_z = 3600$ Hz, passband = 400 Hz, stopband = 60 Hz, $A_{\max} = 0.1$ dB, and $A_{\min} = 40$ dB. The circuit must have provision for frequency tuning of its individual stages.

Solution. The aforementioned FILDES program indicates that we need a sixth-order filter with the following individual-stage parameters:

$$f_{01} = 3460.05 \text{ Hz} \qquad f_{z1} = 3600 \text{ Hz} \qquad Q_1 = 31.4$$

$$f_{02} = 3745.0 \text{ Hz} \qquad f_{z2} = 3600 \text{ Hz} \qquad Q_2 = 31.4$$

$$f_{03} = 3600.0 \text{ Hz} \qquad f_{z3} = 3600 \text{ Hz} \qquad Q_3 = 8.72$$

Moreover, the actual stopband attenuation is 45 dB. This filter is readily designed using three biquad sections, namely, a high-pass notch, followed by a low-pass notch, followed by a symmetric notch (see Problem 4.13).

4.3 GENERALIZED IMPEDANCE CONVERTERS

Impedance converters are active *RC* circuits designed to simulate frequency-dependent elements such as inductances for use in active filter synthesis. Among the various configurations, one that has gained prominence is the *generalized impedance converter* (GIC) of Fig. 4.13, which can be used not only to simulate inductances, but also to synthesize frequency-dependent resistances.

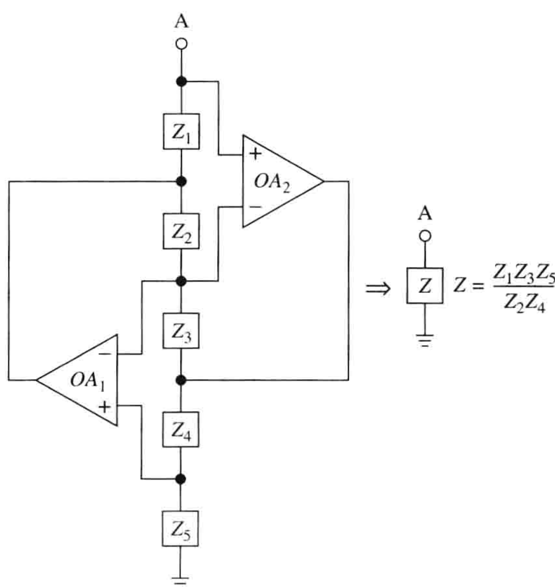


FIGURE 4.13
Generalized impedance converter (GIC).

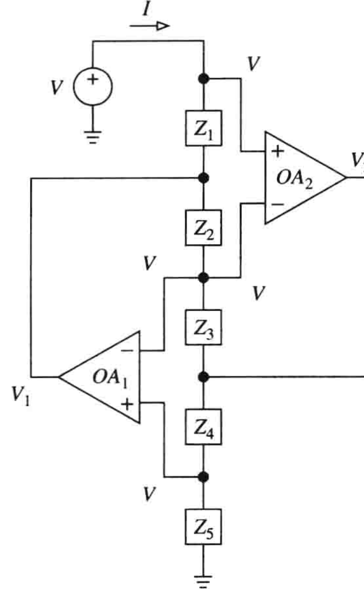


FIGURE 4.14
Finding the equivalent impedance of a GIC toward ground.

To find the equivalent impedance Z seen looking into node A, we apply a test voltage V as in Fig. 4.14, we find the resulting current I, and then we let $Z = V/I$. Exploiting the fact that each op amp keeps $V_n = V_p$, we have labeled the voltages at the input nodes of both op amps as V. By Ohm's law, we have

$$I = \frac{V - V_1}{Z_1}$$

Summing currents at the node common to Z_2 and Z_3 and at the node common to Z_4 and Z_5 we obtain, respectively,

$$\frac{V_1 - V}{Z_2} + \frac{V_2 - V}{Z_3} = 0 \qquad \frac{V_2 - V}{Z_4} + \frac{0 - V}{Z_5} = 0$$

Eliminating V_1 and V_2, and solving for the ratio $Z = V/I$, we get

$$Z = \frac{Z_1 Z_3 Z_5}{Z_2 Z_4} \tag{4.10}$$

Depending on the type of components we use for Z_1 through Z_5, we can configure the circuit for various impedance types. The most interesting and useful ones are as follows:

1. All Zs are resistances, except Z_2 (or Z_4), which is a capacitance. Letting $Z_2 = 1/j\omega C_2$ in Eq. (4.10) gives

$$Z = \frac{R_1 R_3 R_5}{(1/j\omega C_2) R_4} = j\omega L \tag{4.11a}$$

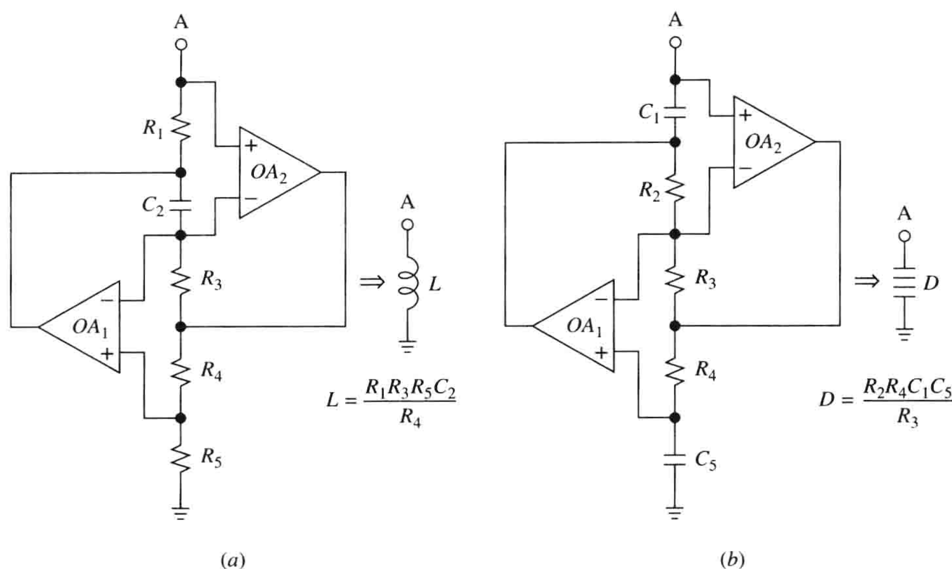


FIGURE 4.15
(*a*) Inductance simulator and (*b*) *D*-element realization.

$$L = \frac{R_1 R_3 R_5 C_2}{R_4} \tag{4.11b}$$

indicating that the circuit simulates a *grounded inductance.* This is depicted in Fig. 4.15*a*. If desired, this inductance can be adjusted by varying one of the resistances, say, R_5.

2. All *Z*s are resistances, except for Z_1 and Z_5, which are capacitances. Letting $Z_1 = 1/j\omega C_1$ and $Z_5 = 1/j\omega C_5$ in Eq. (4.10) gives

$$Z = \frac{(1/j\omega C_1)R_3(1/j\omega C_5)}{R_2 R_4} = -\frac{1}{\omega^2 D} \tag{4.12a}$$

$$D = \frac{R_2 R_4 C_1 C_5}{R_3} \tag{4.12b}$$

The circuit now simulates a *grounded frequency-dependent negative resistance* (grounded FDNR). Since a capacitance produces a voltage proportional to the integral of the current, the FDNR (or *D element,* as it is often called) can be viewed as an element that integrates current twice. Its GIC realization and circuit symbol are shown in Fig. 4.15*b*, and its application will be illustrated shortly. The *D* element can be adjusted by varying one of the resistances.

Figure 4.16 shows another popular realization of the *D* element (see Problem 4.19). Needless to say, the simulated impedances can be no better than the resistances, capacitances, and op amps utilized in their simulation. For good results, use metal-film resistors and NPO ceramic capacitors for temperature stability and polypropylene capacitors for high-*Q* performance. And use a dual op amp with sufficiently fast dynamics (see Section 6.6).

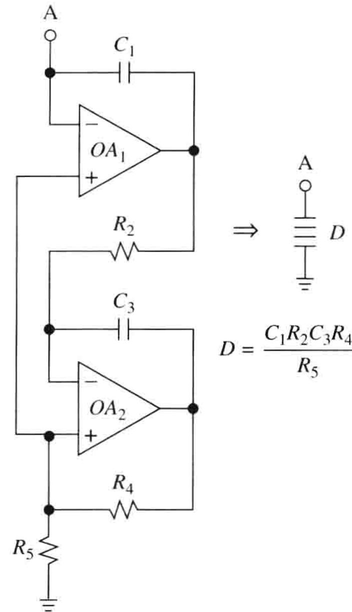


FIGURE 4.16
Alternative D-element realization.

Synthesis Using Grounded Inductances

A popular GIC application is the realization of inductorless filters starting from passive *RLC* filter prototypes. To this end we first design an *RLC* filter meeting the given specifications; then we replace its inductances with synthetic inductances realized with GICs. Note, however, that this direct one-to-one replacement is applicable only if the inductances in the prototype are of the grounded type.

A classic example is offered by the band-pass prototype of Fig. 4.17*a*. This is a band-pass filter because low-frequency signals are shunted by L, high-frequency signals are shunted by C, and intermediate-frequency signals are passed because of resonance. Once the filter specifications are known, we first find a set of *RLC* values meeting the specification; then we replace the original inductance with a GIC inductance simulator to end up with a circuit containing only resistances and capacitances. The result is the *dual-amplifier band-pass* (DABP) filter of Fig. 4.17*b*.

EXAMPLE 4.8. In the circuits of Fig. 4.17 specify component values for a band-pass response with $f_0 = 100$ kHz and $Q = 25$.

Solution. The *RLC* prototype gives $V_o/V_i = (Z_C \parallel Z_L)/(R + Z_C \parallel Z_L)$, $Z_C = 1/(j\omega C)$, $Z_L = j\omega L$. Expanding and collecting gives $V_o/V_i = H_{\text{BP}}$, with

$$\omega_0 = 1/\sqrt{LC} \qquad Q = R\sqrt{C/L}$$

Let $C = 1.0$ nF, so $L = 1/[(2\pi f_0)^2 C] = 1/[(2\pi 10^5)^2 \times 10^{-9}] = 2.533$ mH and $R = Q(L/C)^{1/2} = 25(2.533 \times 10^{-3}/10^{-9})^{1/2} = 39.79$ kΩ. Next, specify the components for the GIC. To simplify inventory, use equal capacitances and equal resistances. Thus, $C_2 = 1.0$ nF. Moreover, using Eq. (4.11*b*), we get $R_1 = R_3 = R_4 = R_5 = (L/C_2)^{1/2} = (2.533 \times 10^{-3}/10^{-9})^{1/2} = 1.592$ kΩ.

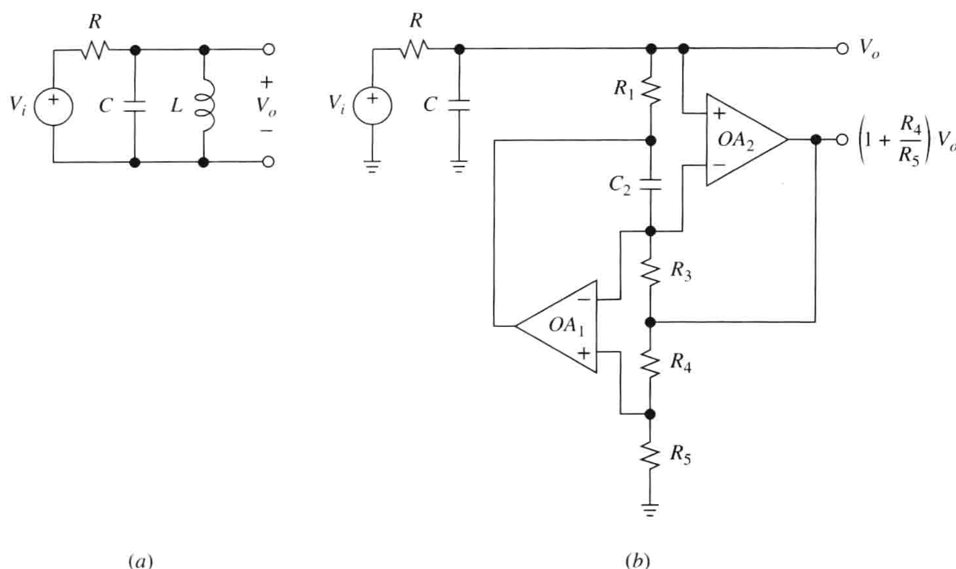


FIGURE 4.17
(*a*) Passive band-pass filter prototype and (*b*) active realization using an inductance simulator.

We observe that the node designated as V_o in Fig. 4.17*b* is prone to external loading. This can be avoided by using the response from the low-impedance output of OA_2, where it is available with a gain of $1 + R_4/R_5$. With equal resistances, this gain is 2 V/V. If unity gain is desired, replace R with a voltage divider, in the manner of Example 3.9.

Using L as given in Eq. (4.11*b*), we have $\omega_0 = \sqrt{R_4/R_1R_3R_5C_2C}$ and $Q = R\sqrt{R_4C/R_1R_3R_5C_2}$, so the sensitivities are

$$S_{R_1}^{\omega_0} = S_{C_2}^{\omega_0} = S_{R_3}^{\omega_0} = -S_{R_4}^{\omega_0} = S_{R_5}^{\omega_0} = S_{C}^{\omega_0} = -1/2$$

$$S_R^Q = 1 \qquad S_{R_1}^Q = S_{C_2}^Q = S_{R_3}^Q = -S_{R_4}^Q = S_{R_5}^Q = -S_C^Q = -1/2$$

These fairly low values are typical of filters based on the ladder simulation approach. If the circuit is implemented with $C_2 = C$ and $R_5 = R_4 = R_3 = R_1$, then $\omega_0 = 1/RC$ and $Q = R/R_1$. This resistance spread compares quite favorably with that of the multiple-feedback band-pass filter, which is $4Q^2$. Moreover, the DABP filter is easily tuned since R_1 (or R_3) adjusts ω_0, and R adjusts Q. Even though the circuit uses two op amps instead of one, it has been proved[6] that if their open-loop frequency characteristics are matched, as is usually the case with dual packages, the op amps tend to compensate for each other's deficiencies, resulting in fairly small deviations of Q and ω_0 from their design values. Owing to these advantages, the DABP filter is a highly recommended configuration.

Synthesis Using FDNRs

As an example of active filter synthesis using FDNRs, consider the *RLC* filter of Fig. 4.18*a*. Low-frequency signals make L a short circuit and C an open, so these

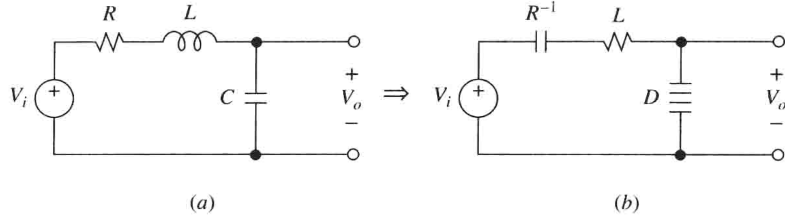


FIGURE 4.18
Low-pass *RLC* filter prototype and its *CRD* equivalent.

signals are passed. High-frequency signals make L an open and C a short, so they are rejected twice, indicating a second-order low-pass response. Since L is not a grounded inductance, we cannot replace it with a simulated one. This obstacle is avoided by the artifice[7] of dividing each element value in the original network by $j\omega$. This transforms resistances into capacitances, inductances into resistances, and capacitances into D elements as

$$\frac{R}{j\omega} \rightarrow \frac{1}{j\omega R^{-1}} \qquad \text{(capacitance of value } R^{-1}\text{)} \tag{4.13a}$$

$$\frac{j\omega L}{j\omega} \rightarrow L \qquad \text{(resistance of value } L\text{)} \tag{4.13b}$$

$$\frac{1/j\omega C}{j\omega} \rightarrow -\frac{1}{\omega^2 C} \qquad (D \text{ element of value } C) \tag{4.13c}$$

The transformed network is shown in Fig. 4.18*b*. It can be proven[3] that dividing all the impedances of a network by the same factor yields a modified network with the same transfer function as the original one. Consequently, the modified circuit of Fig. 4.18*b* not only retains the original response, but is also realizable with a GIC since the transformation has eliminated the floating inductance while creating a grounded D element, which is amenable to GIC simulation.

EXAMPLE 4.9. Using the *RLC* circuit of Fig. 4.18*a* as a prototype, design a GIC low-pass filter with $f_0 = 1$ kHz and $Q = 5$.

Solution. The transformed circuit of Fig 4.18*b* gives, by the voltage divider formula, $V_o/V_i = (-1/\omega^2 C)/(1/j\omega R^{-1} + L - 1/\omega^2 C) = 1/(1 - \omega^2 LC + j\omega RC) = H_{\text{LP}}$, where

$$\omega_0 = 1/\sqrt{LC} \qquad Q = \sqrt{L/C}/R$$

Let the capacitance denoted as R^{-1} be 100 nF. Since $Q\omega_0 = 1/RC$, the value of the D element is $R^{-1}/Q\omega_0 = (100 \times 10^{-9})/(5 \times 2\pi \times 10^3) = 10^{-11}/\pi \text{ s}^2/\Omega$. Finally, the resistance denoted as L is $1/\omega_0^2 C = 1/[(2\pi \times 10^3)^2 \times 10^{-11}/\pi] = 7.958$ kΩ (use 8.06 kΩ, 1%).

Next, specify the components of the GIC, using equal components to simplify inventory. Let $C_1 = C_2 = 10$ nF. By Eq. (4.12*b*), $R_2 = R_3 = R_4 = D/C_2 C_5 = (10^{-11}/\pi)/(10^{-8})^2 = 31.83$ kΩ (use 31.6 kΩ, 1%). The circuit is shown in Fig. 4.19.

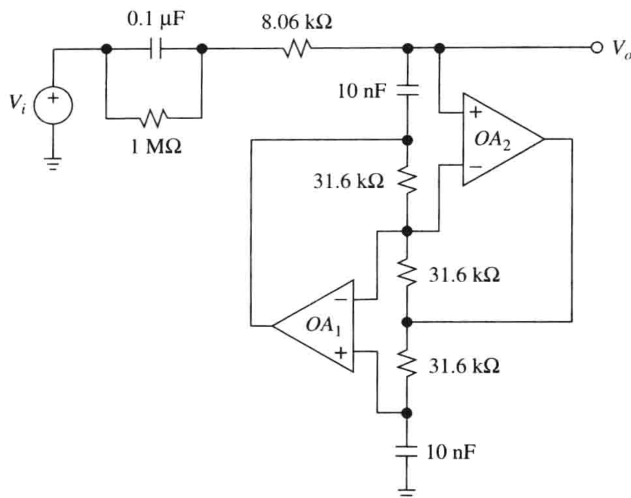


FIGURE 4.19
Low-pass filter using an FDNR.

Remark. In order to provide a dc path for the tiny inverting-input bias current of OA_2, a resistive termination is required. This is performed by the 1-MΩ resistance, whose large value will have little effect on filter performance over the frequency range of interest. A good choice for the op amps is a FET-input dual op amp. To avoid output loading, a buffer can be used.

4.4 DIRECT DESIGN

The interstage isolation properties of cascaded filters, while desirable from the viewpoint of modularity, render the overall response particularly sensitive to individual-stage parameter variations stemming from tolerance, thermal drift, and aging. Of special concern are the high-Q stages, where even a small component variation in a single stage may drastically alter the response of the entire cascade. On the other hand, it has long been recognized that *RLC* filters of the doubly terminated ladder type enjoy the lowest sensitivities to component variations. The ladder structure is a tightly coupled system in which sensitivity is spread out over its elements as a group rather than being confined to specific ones. Sensitivity considerations, together with the wealth of knowledge available in the area of passive *RLC* network synthesis, provide the motivation for the ladder simulation approach.

The starting point is a passive *RLC* ladder prototype, which is designed using suitable filter tables or computer programs. The filter is then realized in active form by replacing its inductors with simulated ones, that is, with active circuits specifically designed to simulate inductance behavior. The resulting active network retains the low-sensitivity advantages of its *RLC* prototype, a feature that makes it suited to applications with stringent specifications.

Figure 4.20 shows the general form of a doubly terminated, series-resonant *RLC* ladder, one of the most frequently used *RLC* prototypes in active filter synthesis. Physically, its behavior is explained as follows. At low frequencies, where the

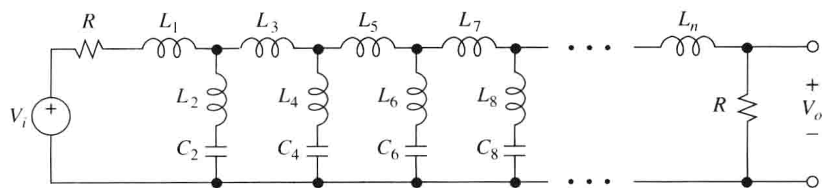


FIGURE 4.20
Doubly terminated series-resonant *RLC* ladder.

inductances act as shorts and the capacitances as opens, the ladder provides a direct signal path from input to output. Low-frequency signals are thus passed, and the dc gain is $R/(R+R) = 1/2$ V/V.

At high frequencies, where the capacitances act as shorts, the ladder becomes predominantly inductive and, as such, it presents considerable impedance to signal propagation. Thus, high-frequency signals are attenuated.

At intermediate frequencies, due to the series resonance of the *LC* elements in each leg, the response exhibits a series of notches, one for each leg. Consequently, the ladder provides a low-pass response with notches, or an *elliptic low-pass response.* The order n of the response is twice the number of legs plus 1, that is, n is odd. If the rightmost inductance is eliminated, then n is decreased by 1 and becomes even. Suppressing the inductances in the ladder legs eliminates the resonances and therefore the stopband notches. This reduced ladder version, referred to as *all-pole ladder,* can then be used to synthesize the Butterworth, Chebyshev, or Bessel responses.

The individual element values are tabulated in filter handbooks[8] or can be calculated by computer.[9] Table 4.2 shows an example of tabulated data. Element values are normalized for a cutoff frequency of 1 rad/s and 1 Ω; however, they are readily adapted to actual frequencies by dividing all reactive elements by the desired cutoff frequency ω_c of the filter.

Low-Pass Filter Design

As is, the ladder of Fig. 4.20 is not amenable to GIC simulation because it contains floating inductances. This obstacle is overcome by applying the $1/j\omega$ transformation discussed in Section 4.3, after which the resistances are changed to capacitances, the inductances to resistances, and the capacitances to D elements. The resulting *CRD* structure is then simulated with grounded FDNRs.

In addition to applying the $1/j\omega$ transformation, we must also frequency-scale the normalized ladder elements to achieve the desired cutoff frequency, and impedance-scale the resulting elements to obtain practical values in the final circuit. The three steps can be carried out at once via the following transformations:[3]

$$C_{\text{new}} = 1/k_z R_{\text{old}} \tag{4.14a}$$

$$R_{j(\text{new})} = (k_z/\omega_c)L_{j(\text{old})} \tag{4.14b}$$

$$D_{j(\text{new})} = (1/k_z\omega_c)C_{j(\text{old})} \tag{4.14c}$$

where $j = 1, 2, \ldots, n$. Here the element values of the *RLC* prototype are referred to as *old,* those of the transformed *RCD* network as *new,* ω_c is the desired cutoff

TABLE 4.2
Element values for doubly terminated Butterworth and Chebyshev low-pass filters

Butterworth low-pass element values (1-rad/s bandwidth)

n	L_1	C_2	L_3	C_4	L_5	C_6	L_7	C_8	L_9	C_{10}
2	1.414	1.414								
3	1.000	2.000	1.000							
4	0.7654	1.848	1.848	0.7654						
5	0.6180	1.618	2.000	1.618	0.6180					
6	0.5176	1.414	1.932	1.932	1.414	0.5176				
7	0.4450	1.247	1.802	2.000	1.802	1.247	0.4450			
8	0.3902	1.111	1.663	1.962	1.962	1.663	1.111	0.3902		
9	0.3473	1.000	1.532	1.879	2.000	1.879	1.532	1.000	0.3473	
10	0.3129	0.9080	1.414	1.782	1.975	1.975	1.782	1.414	0.9080	0.3129

Chebyshev low-pass element values (1-rad/s bandwidth)

n	L_1	C_2	L_3	C_4	L_5	C_6	L_7	C_8	R_2
					0.1-dB ripple				
2	0.84304	0.62201							0.73781
3	1.03156	1.14740	1.03156						1.00000
4	1.10879	1.30618	1.77035	0.81807					0.73781
5	1.14681	1.37121	1.97500	1.37121	1.14681				1.00000
6	1.16811	1.40397	2.05621	1.51709	1.90280	0.86184			0.73781
7	1.18118	1.42281	2.09667	1.57340	2.09667	1.42281	1.18118		1.00000
8	1.18975	1.43465	2.11990	1.60101	2.16995	1.58408	1.94447	0.87781	0.73781
					0.5-dB ripple				
3	1.5963	1.0967	1.5963						1.0000
5	1.7058	1.2296	2.5408	1.2296	1.7058				1.0000
7	1.7373	1.2582	2.6383	1.3443	2.6383	1.2582	1.7373		1.0000
					1.0-dB ripple				
3	2.0236	0.9941	2.0236						1.0000
5	2.1349	1.0911	3.0009	1.0911	2.1349				1.0000
7	2.1666	1.1115	3.0936	1.1735	3.0936	1.1115	2.1666		1.0000

frequency, and k_z is an appropriate impedance-scaling factor to be chosen on the basis of the desired impedance levels in the final circuit.

EXAMPLE 4.10. Figure 4.21 (top) shows a ladder prototype suitable for the GIC realization of a sharp-cutoff smoothing filter for audio D/A converters.[10] The ladder provides a seventh-order Cauer low-pass response with $A_{max} = 0.28$ dB and $A_{min} = 60$ dB at $f_s = 1.252 f_c$. Design an FDNR implementation with $f_c = 15$ kHz.

Solution. First convert the normalized *RLC* prototype to a *CRD* network. Let us arbitrarily decide to use 1-nF capacitances throughout. Since the 1-Ω resistances must change to 1-nF capacitances, Eq. (4.14*a*) gives $k_z = 1/10^{-9} = 10^9$.

By Eq. (4.14*b*), $R_{1(\text{new})} = L_{1(\text{old})} \times 10^9/(2\pi \times 15 \times 10^3) = 1.367 \times 10{,}610 = 14.5$ kΩ, and $R_{2(\text{new})} = 0.1449 \times 10{,}610 = 1.54$ kΩ; by Eq. (4.14*c*), $D_{2(\text{new})} = C_{2(\text{old})}/(10^9 \times 2\pi \times 15 \times 10^3) = 1.207 \times 1.061 \times 10^{-14} = 1.281 \times 10^{-14}\ \text{s}^2/\Omega$. Applying similar transformations to the other elements, we end up with the *CRD* network of Fig. 4.21 (center).

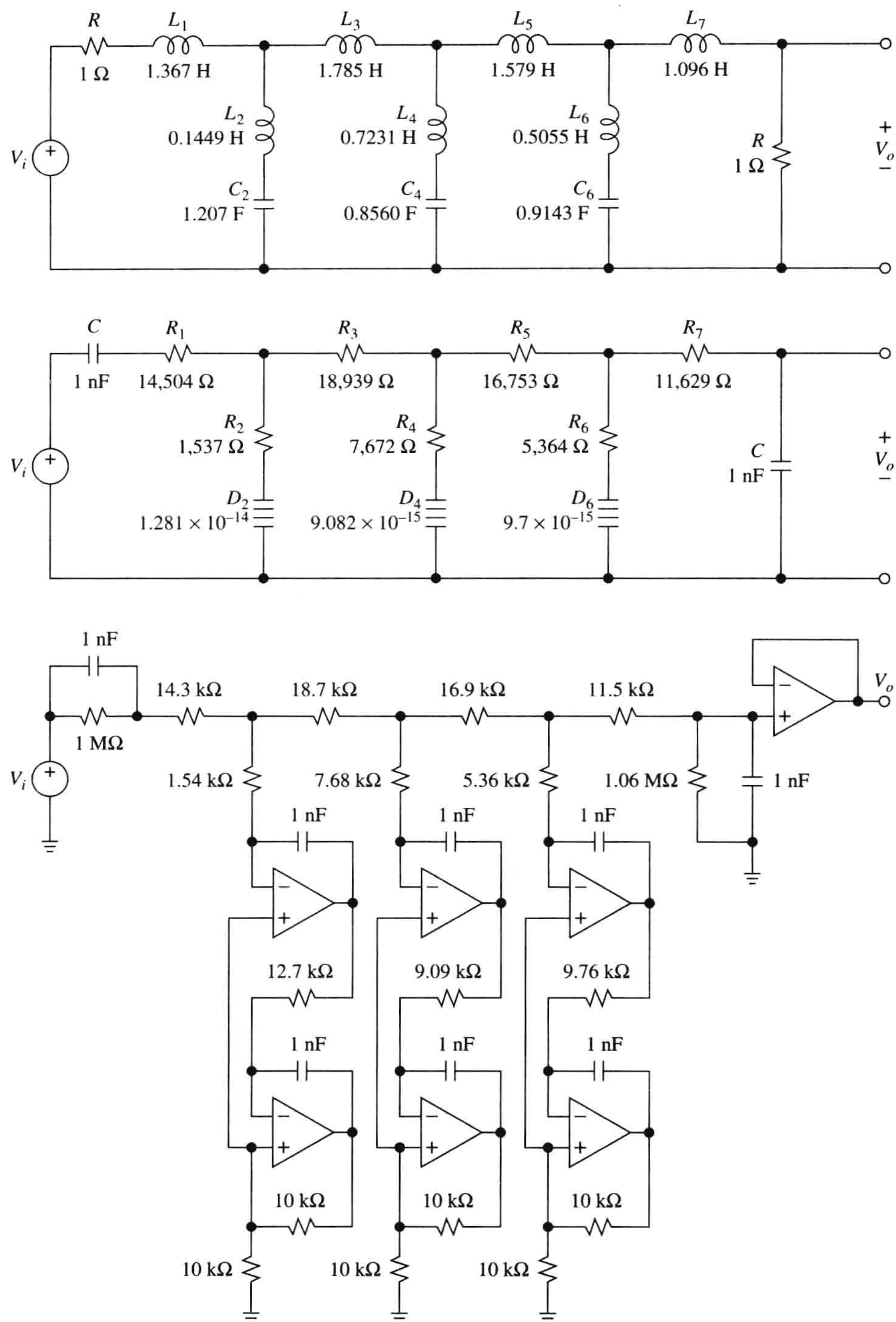


FIGURE 4.21
Seventh-order 0.28/60-dB elliptic low-pass filter of Example 4.10. Top: normalized *RLC* prototype; center: *CRD* equivalent, with *D*-element values in square seconds per ohm; and bottom: active realization using FDNRs.

Finally, we find the elements in the FDNRs. Let us use the FDNRs of Fig. 4.16 with $R_4 = R_5 = 10$ kΩ. Then, Eq. (4.12*b*) gives, for the leftmost FDNR, $R_2 = D/C^2 = 1.281 \times 10^{-14}/(10^{-9})^2 = 12.81$ kΩ (use 12.7 kΩ, 1%). We similarly calculate the remaining FDNRs and end up with the realization of Fig. 4.21 (bottom), where the resistances have been rounded off to 1% standard values.

Note again the use of the 1-MΩ resistance at the input end to provide a dc path for the op amps. To ensure a dc gain of $\frac{1}{2}$ V/V, this resistance must be counterbalanced by a 1.061-MΩ resistance at the output. To avoid loading problems, an output buffer is used. The FDNRs can be implemented with dual FET-input op amps. If desired, each FDNR can be tuned by adjusting one of its resistances.

High-Pass Filter Design

The ladder network of Fig. 4.20, though of the low-pass type, can also serve as a prototype for high-pass filters provided we replace the inductances with capacitances, the capacitances with inductances, and use reciprocal element values to maintain frequency normalization at 1 rad/s. The transformed network provides a response with characteristics reciprocal to the original one, that is, a Cauer high-pass response with a cutoff frequency of 1 rad/s and with notches located at reciprocal positions of the low-pass prototype. Suppressing the capacitances in the legs of the transformed ladder eliminates the stopband notches. This reduced ladder can then be used to synthesize the Butterworth, Chebyshev, or Bessel responses.

In either case, the inductances of the transformed ladder are of the grounded type and as such can be simulated with GICs. After the low-pass to high-pass transformation, the elements must be frequency-scaled to the desired cutoff frequency and impedance-scaled to practical impedance levels. The three steps can be carried out at once via the following transformations:[3]

$$R_{\text{new}} = k_z/R_{\text{old}} \tag{4.15a}$$

$$C_{j(\text{new})} = 1/(k_z \omega_c L_{j(\text{old})}) \tag{4.15b}$$

$$L_{j(\text{new})} = k_z/(\omega_c C_{j(\text{old})}) \tag{4.15c}$$

where the meaning of the notation is similar to Eq. (4.14).

EXAMPLE 4.11. Design an elliptic high-pass filter with $f_c = 300$ Hz, $f_s = 150$ Hz, $A_{\max} = 0.1$ dB, and $A_{\min} = 40$ dB.

Solution. Using standard filter tables[8] or filter-design computer programs,[9] it is found that the specifications can be met with a fifth-order filter whose low-pass prototype has the element values of Fig. 4.22 (top). The actual attenuation at the edge of the stopband is $A(f_s) = 43.4$ dB.

Let us arbitrarily impose $R_{\text{new}} = 100$ kΩ, so $k_z = 10^5$, by Eq. (4.15*a*). Using Eq. (4.15*b*), $C_{1(\text{new})} = 1/(10^5 \times 2\pi \times 300 \times 1.02789) = 5.161$ nF. Using Eq. (4.15*c*), $L_{2(\text{new})} = 10^5/(2\pi \times 300 \times 1.21517) = 43.658$ H. Applying similar transformations to the other elements, we end up with the high-pass ladder of Fig. 4.22 (center).

Finally, we find the elements in the GICs. Let $C = 10$ nF and impose equal resistances. Then, Eq. (4.11) requires, for the leftmost GIC, $R_1 = R_3 = R_4 = R_5 = \sqrt{L/C} = \sqrt{43.658/10^{-8}} = 66.07$ kΩ. Likewise, the resistances for the other GIC are found to be 75.32 kΩ. The final circuit is shown in Fig. 4.22 (bottom), where the resistances have been rounded off to 1% standard values. To avoid output loading, a voltage buffer can be used.

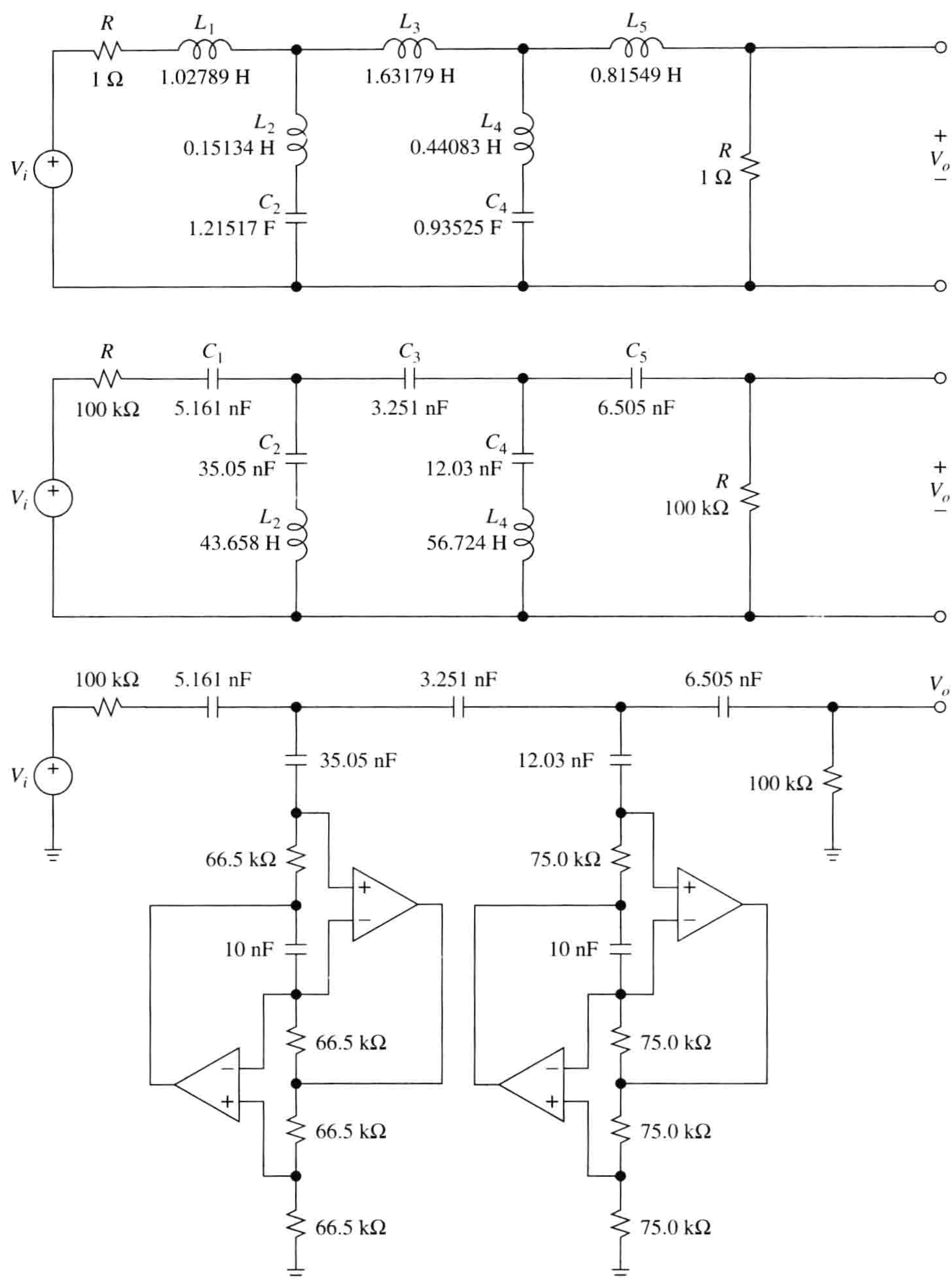


FIGURE 4.22
Fifth-order 0.1/40-dB elliptic high-pass filter of Example 4.11. Top: normalized *RLC* prototype; center: high-pass equivalent; and bottom: active realization using simulated inductances.

4.5
THE SWITCHED CAPACITOR

The filters investigated so far, known as *continuous-time filters,* are characterized by the fact that H_0 and Q are usually controlled by component ratios and ω_0 is controlled by component products. Though ratios can easily be maintained with temperature and time by using devices with adequate tracking capabilities, products are inherently more difficult to control. Moreover, IC processes do not lend themselves to the fabrication of resistances and capacitances with the magnitudes (10^3 to $10^6\ \Omega$ and 10^{-9} to 10^{-6} F) and accuracies (1% or better) typically required in audio and instrumentation applications.

If filter functions are to coexist with digital functions on the same chip, filters must be realized with the components that are most natural to VLSI processes, namely, MOS transistors and small MOS capacitors. This constraint has led to the development of switched-capacitor (SC) filters,[11–13] which simulate resistors by periodically operating MOS capacitors with MOSFET switches, and produce time constants that depend on capacitance ratios rather than *R-C* products.

To illustrate, let us start with the basic MOSFET-capacitor arrangement of Fig. 4.23*a*. The transistors are *n*-channel enhancement types, characterized by a low channel resistance (typically $<10^3\ \Omega$) when the gate voltage is high, and a high resistance (typically $>10^{12}\ \Omega$) when the gate voltage is low. With an off/on ratio this high, a MOSFET can be regarded for all practical purposes as a switch. If the gates are driven with nonoverlapping out-of-phase clock signals of the type in Fig. 4.23*b*, the transistors will conduct on alternate half cycles, thus providing a single-pole double-throw (SPDT) switch function with break-before-make characteristics.

Referring to the symbolic switch representation of Fig. 4.24*a* and assuming $V_1 > V_2$, we observe that flipping the switch to the left charges C to V_1, and flipping it to the right discharges C to V_2. The net charge transfer from V_1 to V_2 is $\Delta Q = C(V_1 - V_2)$. If the switch is flipped back and forth at a rate of f_{CK} cycles per second, the charge transferred in 1 second from V_1 to V_2 defines an average current $I_{avg} = f_{CK} \times \Delta Q$, or

$$I_{avg} = Cf_{CK}(V_1 - V_2) \tag{4.16}$$

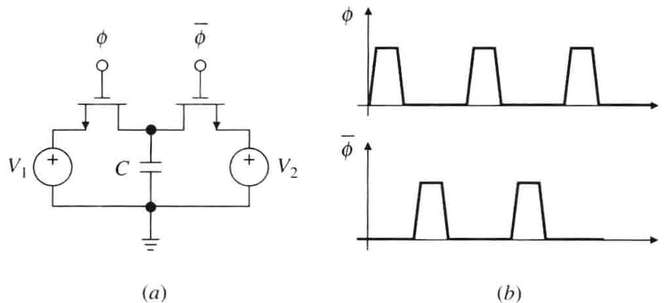


FIGURE 4.23
Switched capacitor using a MOSFET SPDT switch, and clock drive for the MOSFETs.

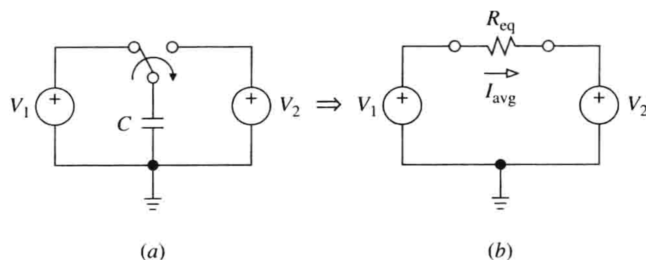


FIGURE 4.24
Resistance simulation using a switched capacitor.

Note that charge is flowing in packets rather than continuously. However, if f_{CK} is made sufficiently higher than the highest-frequency components of V_1 and V_2, the process can be regarded as continuous, and the switch-capacitor combination can be modeled with an equivalent resistance

$$R_{eq} = \frac{V_1 - V_2}{I_{avg}} = \frac{1}{Cf_{CK}} \tag{4.17}$$

The model is depicted in Fig. 4.24*b*. Let us investigate how such a resistance can be used to implement, what by now has proved to be the workhorse of active filters, namely, the integrator.

SC Integrators

As we know, the *RC* integrator of Fig. 4.25*a* yields $H(j\omega) = -1/(j\omega/\omega_0)$, where the unity-gain frequency is given by

$$\omega_0 = \frac{1}{R_1C_2} \tag{4.18}$$

Replacing R_1 by an SC resistance gives the SC integrator of Fig. 4.25*b*. If the input frequency ω is such that

$$\omega \ll \omega_{CK} \tag{4.19}$$

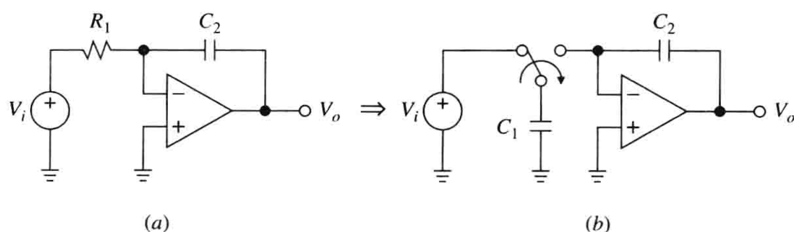


FIGURE 4.25
Converting an *RC* integrator to an SC integrator.

where $\omega_{CK} = 2\pi f_{CK}$, then current flow from V_i to the summing node can be regarded as continuous, and ω_0 is found by substituting R_{eq} into Eq. (4.18),

$$\omega_0 = \frac{C_1}{C_2} f_{CK} \tag{4.20}$$

This expression reveals three important features that hold for SC filters in general, not just for SC integrators:

1. There are no resistors. This is highly desirable from the viewpoint of IC fabrication, since monolithic resistors are plagued by large tolerances and thermal drift, and also take up precious chip area. Switches, on the other hand, are implemented with MOSFETs, which are the basic ingredients of VLSI technology and occupy very little chip area.
2. The characteristic frequency ω_0 depends on capacitance ratios, which are much easier to control and maintain with temperature and time than *R-C* products. With present technology, ratio tolerances as low as 0.1% are readily achievable.
3. The characteristic frequency ω_0 is proportional to the clock frequency f_{CK}, indicating that SC filters are inherently of the programmable type. Varying f_{CK} will shift the response up or down the frequency spectrum. If, on the other hand, a fixed and stable characteristic frequency is desired, f_{CK} can be generated with a quartz crystal oscillator.

Equation (4.20) also shows that by judicious choice of the values of f_{CK} and the C_1/C_2 ratio, it is possible to avoid undesirably large capacitances even when low values of ω_0 are desired. For instance, with $f_{CK} = 1$ kHz, $C_1 = 1$ pF, and $C_2 = 15.9$ pF, the SC integrator gives $f_0 = (1/2\pi)(1/15.9)10^3 = 10$ Hz. An *RC* integrator with the same f_0 could be implemented, for instance, with $R_1 = 1.59$ MΩ and $C_2 = 10$ nF. Fabricating these components monolithically and maintaining the value of their product within 0.1% would be unrealistic. Current SC filters use capacitances in the range of 0.1 pF to 100 pF, with the 1-pF to 10-pF range being the most common. The upper limit is dictated by die area considerations, and the lower limit by parasitic capacitances of the SC structure.

To minimize the effect of parasitic capacitances and also increase circuit versatility, practical SC integrators are implemented with SPDT switch pairs, in the manner of Fig. 4.26. In Fig. 4.26*a*, flipping the switches down discharges C_1 to zero, and flipping the switches up charges C_1 to V_i. Current will thus flow into the summing junction of the op amp if $V_i > 0$, and out if $V_i < 0$, indicating that the integrator is of the inverting type.

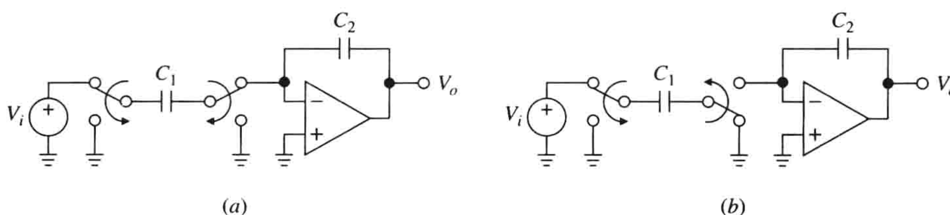


FIGURE 4.26
Inverting and noninverting SC integrators.

Changing the phase of one of the switches yields the circuit of Fig. 4.26*b*. With the switches in the positions shown, the left plate of C_1 is at V_i and the right plate at 0 V. Commutating both switches will discharge C_1 to 0 V, thus pulling charge out of the summing junction if $V_i > 0$, and pushing current into the junction if $V_i < 0$. A simple phase rewiring of the two MOSFETs making up one of the switches inverts the direction of I_{avg}, resulting in an SC integrator of the noninverting type. We shall exploit the availability of this type of integrator in the next sections.

Practical Limitations of SC Filters

There are some important limitations that we need be aware of when applying SC filters.[11] First, there are limits on the permissible range of f_{CK}. The upper limit is dictated by the quality of the MOS switches and the speed of the op amps. Taking 10 pF as a typical switched capacitance and 1 kΩ as a typical resistance of a closed MOS switch, we observe that the time constant is on the order of $10^3 \times 10^{-11} = 10$ ns. Considering that to charge a capacitance to within 0.1% of its final voltage takes about seven time constants ($e^{-7} \cong 10^{-3}$), it follows that the minimum time interval between consecutive switch commutations is on the order of 10^2 ns. This also happens to be the typical time it takes for the step response of a MOS op amp to settle within 0.1% of its final value. Consequently, the upper limit for f_{CK} is in the megahertz range.

The lower practical limit for f_{CK} is dictated by the leakage of open MOS switches and the input bias currents of op amps, both of which tend to discharge the capacitors and, hence, to destroy the accumulated information. At room temperature these currents are in the picoampere range. Assuming a maximum acceptable droop of 1 mV across a capacitor of 10 pF, we have $f_{CK} \geq (1\text{ pA})/[(10\text{ pF}) \times (1\text{ mV})] = 10^2$ Hz. In summary, the permissible clock range is typically $10^2\text{ Hz} < f_{CK} < 10^6$ Hz.

The other important limitation of SC filters stems from their discrete-time rather than continuous-time operation. This is evidenced in Fig. 4.27, which shows the

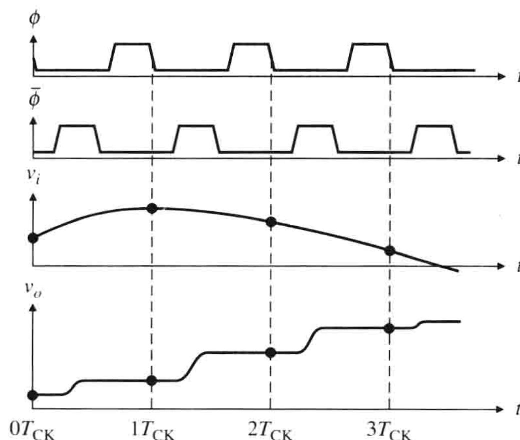


FIGURE 4.27
Noninverting SC integrator waveforms.

input and output waveforms for the noninverting integrator of Fig. 4.26a. Time has been divided into equal intervals according to the clock period T_{CK}. Referring to the actual circuit, we observe that ϕ pulses charge C_1 to v_i, while $\bar{\phi}$ pulses pull the charge accumulated in C_1 out of C_2, causing a step increase in v_o. Because of nonzero switch resistance, this step is gradual.

Letting n denote an arbitrary clock period, we have $v_o[nT_{CK}] = v_o[(n-1)T_{CK}] + \Delta Q[(n-1)T_{CK}]/C_2$, or

$$v_o[nT_{CK}] = v_o[(n-1)T_{CK}] + \frac{C_1}{C_2} v_i[(n-1)T_{CK}] \tag{4.21}$$

where $\Delta Q[(n-1)T_{CK}] = C_1 v_i[(n-1)T_{CK}]$ denotes the charge accumulated by C_1 during the previous ϕ pulse. Equation (4.21) represents a discrete time sequence relating input and output values, which have been emphasized with dots. A well-known Fourier transform property states that delaying a signal by one clock period T_{CK} is equivalent to multiplying its Fourier transform by $\exp(-j\omega T_{CK})$. Taking the Fourier transforms of both sides of Eq. (4.21) gives

$$V_o(j\omega) = V_o(j\omega)e^{-j\omega T_{CK}} + \frac{C_1}{C_2} V_i(j\omega)e^{-j\omega T_{CK}} \tag{4.22}$$

Collecting, solving for the ratio $H(j\omega) = V_o(j\omega)/V_i(j\omega)$, and using Euler's identity $\sin\alpha = (e^{\alpha} - e^{-\alpha})/2j$, we finally obtain the *exact* transfer function of the SC noninverting integrator,

$$H(j\omega) = \frac{1}{j\omega/\omega_0} \times \frac{\pi\omega/\omega_{CK}}{\sin(\pi\omega/\omega_{CK})} \times e^{-j\pi\omega/\omega_{CK}} \tag{4.23}$$

where $\omega_0 = (C_1/C_2)f_{CK}$ and $\omega_{CK} = 2\pi/T_{CK} = 2\pi f_{CK}$.

We observe that in the limit $\omega/\omega_{CK} \to 0$ we obtain the familiar integrator function $H(j\omega) = 1/(j\omega/\omega_0)$, confirming that as long as $\omega_{CK} \gg \omega$, the SC process can be regarded as a continuous-time process. Writing $H(j\omega) = [1/(j\omega/\omega)] \times \epsilon_m \times \exp(-j\epsilon_\phi)$ indicates that in general the SC process introduces a *magnitude error* $\epsilon_m = (\pi\omega/\omega_{CK})/[\sin(\pi\omega/\omega_{CK})]$ and a *phase error* $\epsilon_\phi = -\pi\omega/\omega_{CK}$. The effect of these errors is illustrated in the linear plots of Fig. 4.28 for a noninverting integrator with $\omega_0 = \omega_{CK}/10$.

The ideal magnitude and phase responses are $|H| = 1/(\omega/\omega_0)$ and $\measuredangle H = -90°$. The SC integrator deviation increases with ω until, for $\omega = \omega_{CK}$, the magnitude error becomes infinite and phase undergoes polarity reversal. These results are consistent with well-known sampled-data principles, stating that the effect of sampling a function of time at the rate of f_{CK} samples per second is a replication of its frequency spectrum at integral multiples of f_{CK}.

For $\omega \ll \omega_{CK}$, the effect of the magnitude error is similar to the effect of component tolerance or drift in ordinary RC integrators. As such, it may not be detrimental, especially if the performance requirements are not stringent. To contain this error within tolerable limits, the useful frequency range is limited to a couple of decades below ω_{CK}.

The effect of the phase error, however, is critical since it may cause Q enhancement or even instability. One method of compensating for this error is by alternating the clock phasing of consecutive integrators,[11] as we shall see in Section 4.6.

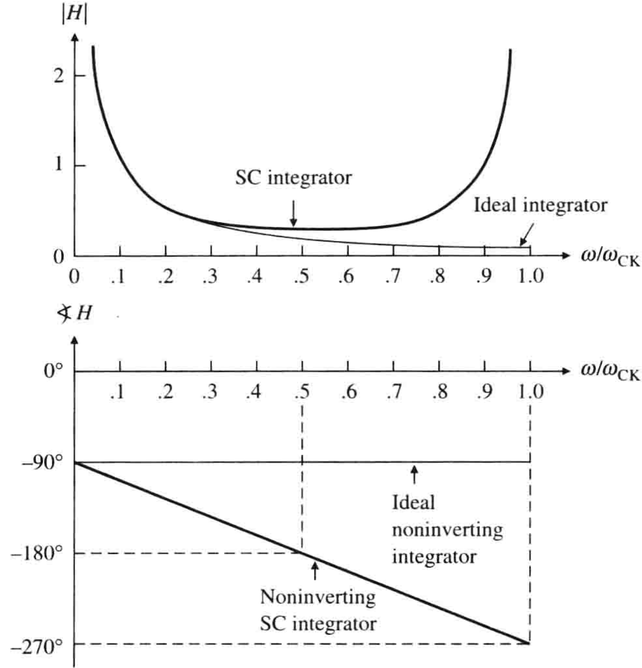


FIGURE 4.28
Magnitude and phase responses of a noninverting SC integrator for the case $\omega_0 = \omega_{CK}/10$.

4.6
SWITCHED-CAPACITOR FILTERS

Switched-capacitor filters are based on the integrator configurations of the previous section. As in the case of continuous-time filters, two popular approaches to SC filter synthesis are the cascade approach and the ladder simulation approach.

Dual-Integrator-Loop Filters

A dual-integrator-loop SC filter can be synthesized by replacing the resistors of a continuous-time prototype with SC equivalents. Figure 4.29 shows the SC implementation of the popular biquad topology of Fig. 3.36. Here OA_2 is a lossless noninverting integrator, a function that requires only one op amp when implemented in SC form. We thus have, for $\omega \ll \omega_{CK}$,

$$V_{LP} = \frac{1}{j\omega/\omega_0} V_{BP}$$

where $\omega_0 = (C_1/C_2) f_{CK}$, by Eq. (4.20). The op amp OA_1 forms a lossy inverting integrator, whose equivalent feedback resistance, simulated by C_3 and the associated switch, sets the value of Q. By Eq. (4.17), this resistance is $R_Q = 1/(C_3 f_{CK})$. With the input switches in the position shown, the leftmost capacitance C_1 is charged to $V_{LP} - V_i$. Flipping the switches down transfers the charge $\Delta Q = C_1(V_{LP} - V_i)$

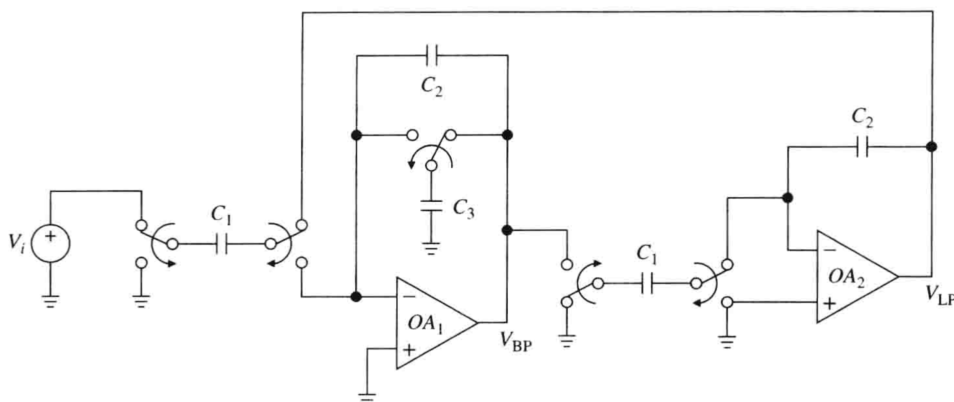


FIGURE 4.29
SC biquad filter.

into the summing junction of OA_1, so the corresponding average current is $I_1 = C_1 f_{CK}(V_{LP} - V_i)$. Summing currents at this junction gives, for $\omega \ll \omega_{CK}$,

$$C_1 f_{CK}(V_{LP} - V_i) + C_3 f_{CK} V_{BP} + j\omega C_2 V_{BP} = 0$$

Substituting $V_{LP} = V_{BP}/(j\omega/\omega_0)$ and collecting gives $V_{BP}/V_i = H_{0BP} H_{BP}$ and $V_{LP}/V_i = H_{0LP} H_{LP}$, where H_{LP} and H_{BP} are the standard second-order low-pass and band-pass responses, and

$$\omega_0 = \frac{C_1}{C_2} f_{CK} \qquad Q = \frac{C_1}{C_3} \qquad H_{0BP} = Q \qquad H_{0LP} = 1 \text{ V/V} \tag{4.24}$$

EXAMPLE 4.12. Assuming $f_{CK} = 100$ kHz in the circuit of Fig. 4.29, specify suitable capacitances for a Butterworth low-pass response with $f_0 = 1$ kHz and a total capacitance of 100 pF or less.

Solution. We have $C_2/C_1 = f_{CK}/(2\pi f_0) = 15.9$ and $C_3/C_1 = 1/Q = \sqrt{2}$. Choose $C_1 = 1$ pF, $C_2 = 15.9$ pF, and $C_3 = 1.41$ pF.

The realization of Fig. 4.29 is by no means unique, nor is it necessarily the best. In fact (see Problem 4.30), its capacitance spread increases with Q to the point of making this arrangement unfeasible. Figure 4.30 shows an SC realization with

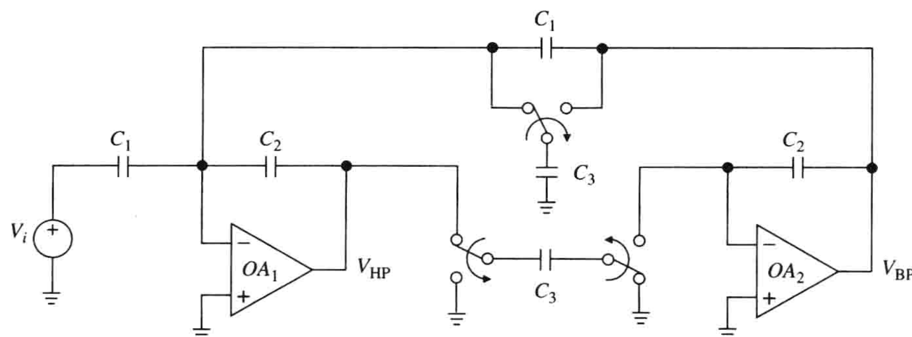


FIGURE 4.30
SC biquad filter with improved capacitance spread.

improved capacitance ratios. The circuit uses an integrator/summer and a noninverting integrator to provide the band-pass and high-pass responses. It can be proved (see Problem 4.31) that

$$\omega_0 = \frac{C_3}{C_2} f_{\text{CK}} \qquad Q = \frac{C_2}{C_1} \qquad H_{0\text{BP}} = -1 \text{ V/V} \qquad H_{0\text{HP}} = \frac{-1}{Q} \tag{4.25}$$

In the next section we investigate the cascade design of higher-order filters using dual-integrator loops, an approach that is particularly attractive when filter specifications are not too stringent. For low-sensitivity applications, the direct synthesis methods discussed next are preferable.

Ladder Simulation

Direct SC filter synthesis uses SC integrators to simulate passive *RLC* ladders. Since it retains the low-sensitivity advantages of ladders, this approach is preferable when filter specifications are more stringent. One of the most frequently used structures is the doubly terminated all-pole ladder of Fig. 4.31, which can be configured for Butterworth, Chebyshev, or Bessel responses, the order n coinciding with the number of reactive elements present. As we know, the required component values are tabulated in filter handbooks or can be calculated by computer.

We observe that the ladder is a repetitive structure of *LC* pairs of the type of Fig. 4.32*a*. The inductance current is

$$I_{k-1} = \frac{V_{k-1} - V_k}{j\omega L_{k-1}}$$

SC integrators are inherently voltage-processing blocks, so to make the above function amenable to SC implementation, we use the artifice of multiplying both sides by a scaling resistance R_s, which converts the current I_{k-1} to a voltage $V'_{k-1} = R_s I_{k-1}$, or

$$V'_{k-1} = \frac{1}{j\omega/\omega_{L_{k-1}}}(V_{k-1} - V_k) \qquad \omega_{L_{k-1}} = \frac{1}{L_{k-1}/R_s}$$

This integration is implemented with an L-integrator of the type also shown in Fig. 4.32*b*. By Eq. (4.20), its capacitances must satisfy $C_0/C_{L_{k-1}} = \omega_{L_{k-1}}$, or

$$C_{L_{k-1}}/C_0 = (L_{k-1}/R_s) f_{\text{CK}} \tag{4.26}$$

Next consider the capacitance C_k, whose voltage is

$$V_k = \frac{1}{j\omega C_k}(I_{k-1} - I_k)$$

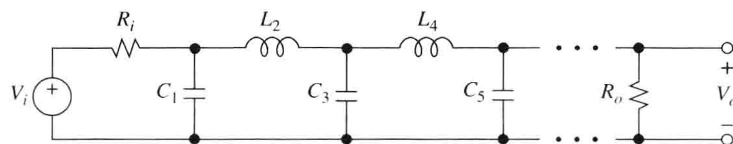


FIGURE 4.31
Doubly terminated all-pole *RLC* ladder.

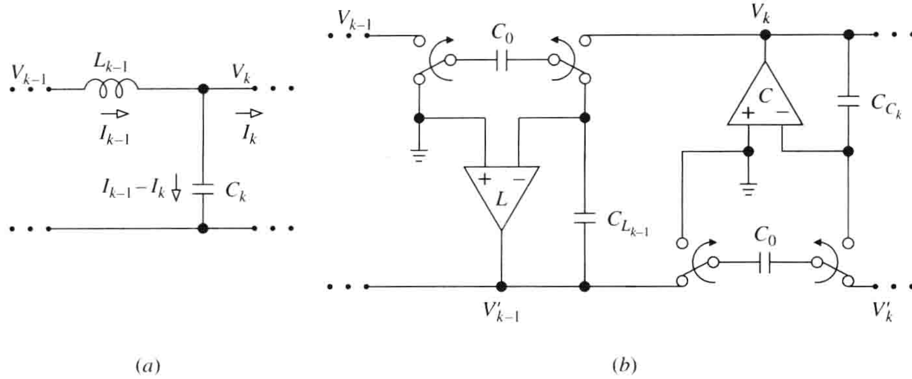


FIGURE 4.32
LC ladder section and its realization in SC form.

Multiplying numerator and denominator by R_s to convert the currents I_{k-1} and I_k to the voltages $V'_{k-1} = R_s I_{k-1}$ and $V'_k = R_s I_k$, we obtain

$$V_k = \frac{1}{j\omega/\omega_{C_k}}(V'_{k-1} - V'_k) \qquad \omega_{C_k} = \frac{1}{R_s C_k}$$

This integration is implemented with a C-integrator of the type also shown in Fig. 4.32*b*. By Eq. (4.20), its capacitances must satisfy $C_0/C_{C_k} = \omega_{C_k}$, or

$$C_{C_k}/C_0 = R_s C_k f_{\text{CK}} \tag{4.27}$$

We thus conclude that if the conditions of Eqs. (4.26) and (4.27) are met, the SC integrators of Fig. 4.32*b* will simulate the *LC* pair of Fig. 4.32*a*. The by-product variables V'_{k-1} and V'_k need not concern us as they are internal to the circuit.

To complete the ladder simulation, we also need SC equivalents of the terminating resistors. This is readily achieved by making the first and last SC integrators of the lossy type. Denoting the capacitances simulating these resistances as C_{R_i} and C_{R_o}, we have

$$C_{R_i}/C_0 = R_i/R_s \qquad C_{R_o}/C_0 = R_o/R_s \tag{4.28}$$

For simplicity we can let $R_i = R_o = R_s = 1\ \Omega$, after which we get $C_{R_i} = C_{R_o} = C_0$.

As an example, Fig. 4.33 shows a fifth-order low-pass SC filter. Since the leftmost reactive element in the ladder prototype is a capacitance, the leftmost integrator is a C-integrator. The rightmost integrator is either a C-integrator or an L-integrator, depending on whether the order n of the filter is odd (as in the example) or even. Moreover, the leftmost and rightmost integrators must be of the lossy type to simulate the terminating resistances. Note also the alternation in the switch phases of adjacent integrators in order to minimize the effects of sampling delays, as mentioned at the end of the previous section.

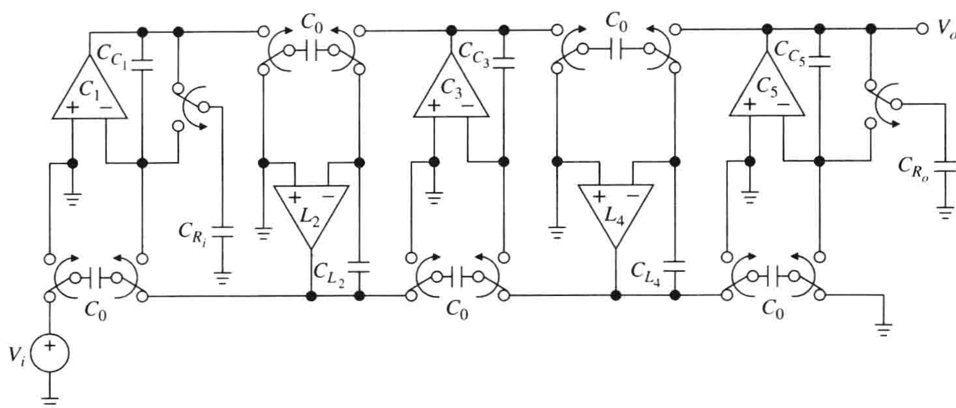


FIGURE 4.33
Fifth-order SC low-pass filter.

Direct Synthesis of Low-Pass Filters

Although the element values of Table 4.2 refer to all-pole ladders with an inductance as the leftmost reactive element, they are readily adapted to ladders with a capacitance as the leftmost reactive element, provided we change the column headings from L_1, $C_2, L_3, C_4, \ldots$ to $C_1, L_2, C_3, L_4, \ldots$. Since the tabulated *RLC* values are normalized for a cutoff frequency of 1 rad/s, they must be frequency-scaled before Eqs. (4.26) and (4.27) can be applied. As discussed in Section 4.4, this requires dividing all reactive values by the cutoff frequency ω_c. Assuming $R_s = 1\ \Omega$, the above equations become

$$C_{C_k}/C_0 = (C_k/\omega_c) f_{\text{CK}} \qquad C_{L_k}/C_0 = (L_k/\omega_c) f_{\text{CK}} \tag{4.29}$$

where C_k and L_k represent the kth normalized reactive element values of the filter prototype.

EXAMPLE 4.13. In the circuit of Fig. 4.33, specify capacitances for a fifth-order Butterworth low-pass response with $f_c = 1$ kHz and $f_{\text{CK}} = 100$ kHz.

Solution. From Table 4.2 we find the following normalized element values: $C_1 = C_5 = 0.618$, $C_3 = 2.000$, and $L_2 = L_4 = 1.618$. Using Eq. (4.29), we obtain $C_{C_1}/C_0 = 0.618 \times 10^5/(2\pi 10^3) = 9.836$, $C_{L_2}/C_0 = 1.618 \times 10^5/(2\pi 10^3) = 25.75$, etc., and $C_{R_i}/C_0 = C_{R_o}/C_0 = 1$. A set of capacitances meeting the above constraints is $C_{R_i} = C_{R_o} = C_0 = 1$ pF, $C_{C_1} = C_{C_5} = 9.84$ pF, $C_{L_2} = C_{L_4} = 25.75$ pF, and $C_{C_3} = 31.83$ pF.

Direct Synthesis of Band-Pass Filters

The low-pass ladder of Fig. 4.31 can also serve as the prototype for other responses. For example, replacing each capacitance by an inductance and vice versa, and using reciprocal element values, the ladder becomes of the high-pass type. Replacing each inductance in the original ladder by a parallel *LC* pair yields a low-pass response with notches, that is, an elliptic low-pass response. Replacing each capacitance in the original ladder by a parallel *LC* pair and each inductance by a series *LC* pair

yields a band-pass response. Replacing each capacitance in the original ladder by a series *LC* pair and each inductance by a parallel *LC* pair yields a band-reject response.

Once the ladder has been transformed, we write circuit equations for each node and branch, and use resistance scaling to convert currents to voltages to render the equations amenable to SC simulation. We shall illustrate the procedure for the band-pass case.

The ladder of Fig. 4.34 (top) is a second-order low-pass prototype. If we replace its capacitance by a parallel *LC* pair and its inductance by a series *LC* pair, we end up with the fourth-order band-pass ladder of Fig. 4.34 (center). *RLC* filter theory states[3] that to achieve a center frequency of 1 rad/s with a normalized bandwidth BW, the element values of the transformed ladder must be related to those of the low-pass prototype as

$$C_{1(\text{new})} = C_{1(\text{old})}/\text{BW} \qquad L_{1(\text{new})} = \text{BW}/C_{1(\text{old})} \tag{4.30a}$$

$$C_{2(\text{new})} = \text{BW}/L_{2(\text{old})} \qquad L_{2(\text{new})} = L_{2(\text{old})}/\text{BW} \tag{4.30b}$$

where the low-pass elements are referred to as *old*, and the band-pass ones as *new*. The former are tabulated in filter handbooks.

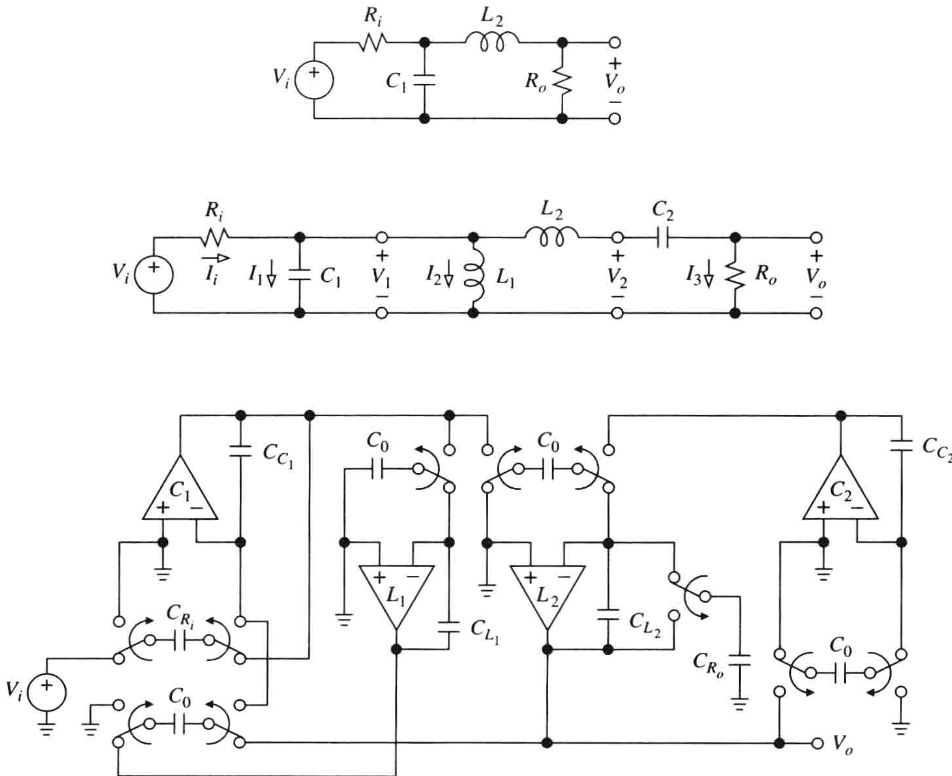


FIGURE 4.34
Fourth-order band-pass filter. Top: second-order *RLC* prototype; center: fourth-order *RLC* equivalent; and bottom: SC realization.

Let us now develop the necessary circuit equations. By KCL, $V_1 = (1/j\omega C_1) \times (I_i - I_2 - I_3)$. Multiplying numerator and denominator by the scaling resistance R_s to convert currents to voltages as $V_i' = R_s I_i$, $V_2' = R_s I_2$, and $V_3' = R_s I_3$, we obtain

$$V_1 = \frac{1}{j\omega/\omega_{C_1}}(V_i' - V_2' - V_3') \qquad \omega_{C_1} = \frac{1}{R_s C_1}$$

By Ohm's law, $I_2 = V_1/j\omega L_1$. Multiplying both sides by R_s gives

$$V_2' = \frac{1}{j\omega/\omega_{L_1}} V_1 \qquad \omega_{L_1} = \frac{1}{L_1/R_s}$$

By Ohm's law, $I_3 = (V_1 - V_2)/j\omega L_2$, or

$$V_3' = \frac{1}{j\omega/\omega_{L_2}}(V_1 - V_2) \qquad \omega_{L_2} = \frac{1}{L_2/R_s}$$

By KVL, $V_2 = V_o + I_3/(j\omega C_2)$, or

$$V_2 = V_o + \frac{1}{j\omega/\omega_{C_2}} V_3' \qquad \omega_{C_2} = \frac{1}{R_s C_2}$$

All equations are realizable with the SC integrators of Section 4.5. An actual implementation is shown in Fig. 4.34 (bottom). The SC capacitance ratios are found via Eq. (4.29) with ω_c replaced by the desired center frequency ω_0.

EXAMPLE 4.14. Specify capacitances in Fig. 4.34 (bottom) for a fourth-order 0.1-dB Chebyshev band-pass response with $f_0 = 1$ kHz, BW $= 600$ Hz, and $f_{CK} = 100$ kHz.

Solution. From Table 4.2 we find the following low-pass prototype element values: $C_1 = 0.84304$ and $L_2 = 0.62201$. The normalized bandwidth is BW $= 600/1000 = 0.6$, so the normalized band-pass ladder elements are $C_1 = 0.84304/0.6 = 1.405$, $L_1 = 0.6/0.84304 = 0.712$, $L_2 = 0.62201/0.6 = 1.037$, and $C_2 = 0.6/0.62201 = 0.9646$.

Using $R_i = R_o = R_s = 1\ \Omega$, and $C_{R_i} = C_{R_o} = C_0 = 1$ pF, we find $C_{C_1} = 10^5 C_1/(2\pi 10^3) = 15.92$, $C_1 = 15.92 \times 1.405 = 22.36$ pF, $C_{L_1} = 15.92 \times 0.712 = 11.33$ pF, $C_{L_2} = 16.51$ pF, and $C_{C_2} = 14.81$ pF.

Switched-capacitor ladder filters are available in a variety of configurations, both in stand-alone form and as part of complex systems such as Codecs. Stand-alone filters are usually preconfigured for commonly used responses, such as the eighth-order Butterworth, Cauer, and Bessel responses provided by the SC filters of the LTC1064 series.

4.7 UNIVERSAL SC FILTERS

Universal SC filters use the dual-integrator-loop configuration to provide the basic second-order responses. These responses can then be cascaded to implement higher-order filters. Two popular and well-documented examples are the LTC1060 and the MF10.

The MF10 Universal SC Filter

The MF10 filter, whose block diagram is shown in Fig. 4.35, consists of two dual-integrator-loop sections, each equipped with an uncommitted op amp to add versatility and facilitate cascading. Each section can be configured independently for the low-pass, band-pass, high-pass, notch, and all-pass responses by means of external resistances. Though these resistances could have been synthesized on-chip using SC techniques, placing them under the control of the user increases the versatility of the circuit. Furthermore, filter parameters are made to depend on resistance ratios, rather than on absolute values, to take advantage of component tracking.

The integrators are of the noninverting type, with the transfer function

$$H(jf) = \frac{1}{jf/f_1} \tag{4.31}$$

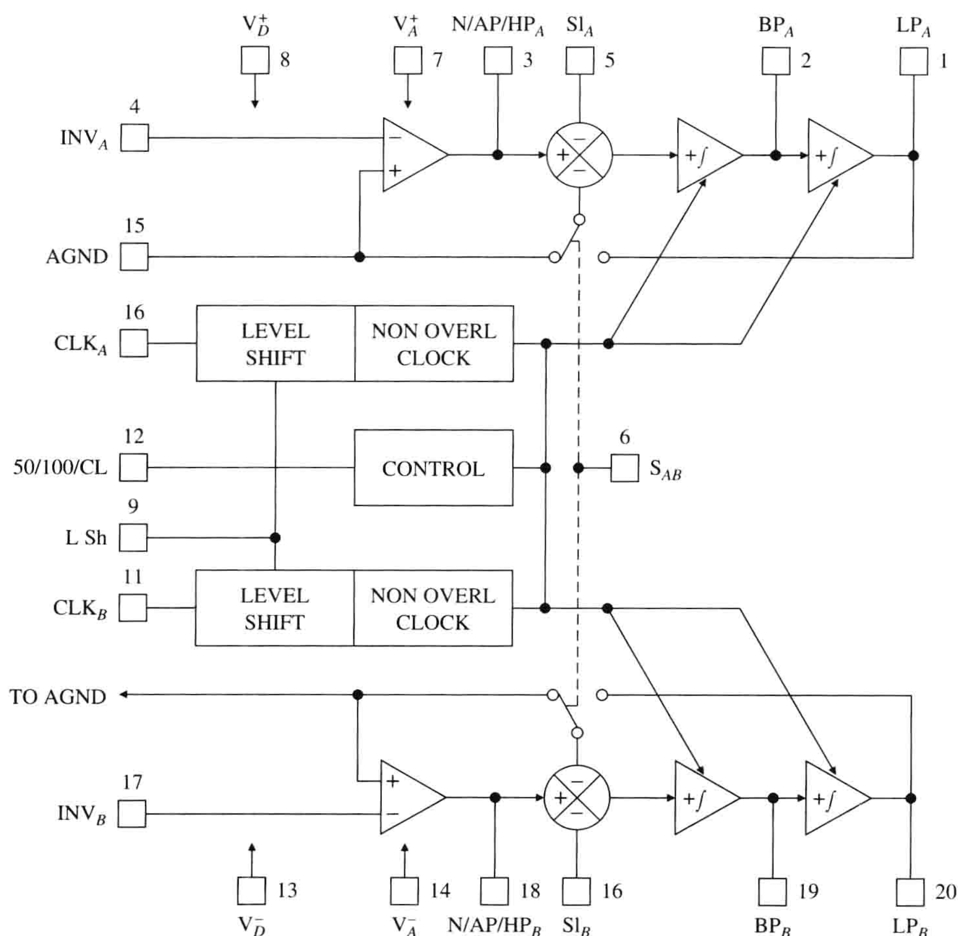


FIGURE 4.35
Block diagram of the MF10 universal monolithic dual SC filter. (Courtesy of Texas Instruments.)

where f_1 is the *integration unity-gain frequency,* and

$$f_1 = \frac{f_{CK}}{100} \quad \text{or} \quad \frac{f_{CK}}{50} \tag{4.32}$$

depending on the voltage level applied at the 50/100/CL frequency-ratio programming pin: tying it to ground enables the 100 ratio, and tying it to the positive supply enables the 50 ratio.

In general, the characteristic frequency f_0 of a section coincides with the unity-gain frequency f_1 of its integrators; however, connecting an external resistance between the LP and INV pins shifts f_0 away from f_1 by an amount controlled by an external resistance ratio. This feature is useful in cascade design, where the resonance frequency of each stage must be set independently while all sections are controlled by the same clock frequency f_{CK}.

For additional flexibility, an internal programming switch is provided, whose position is user-programmable by means of the S_{AB} control pin. Tying this pin to the positive (negative) supply positions the switch to the right (left). Whereas the integrators provide the band-pass and low-pass responses, the input amplifier can be made to provide the high-pass, notch, or all-pass response, depending on the external resistance connections and the internal switch position.

Modes of Operation

Each section can be configured for a variety of different modes. The following are some of the most significant ones; others can be found in the data sheets and application notes.[4]

The circuit of Fig. 4.36 provides the notch, band-pass, and low-pass responses. Since the summing amplifier is outside the two-integrator loop, this mode is faster and allows for a wider range of operating frequencies. Assuming $f \ll f_{CK}$, we have

$$V_N = -\frac{R_2}{R_1}V_i - \frac{R_2}{R_3}V_{BP} \qquad V_{BP} = \frac{V_N - V_{LP}}{jf/f_1} \qquad V_{LP} = \frac{V_{BP}}{jf/f_1}$$

where f_1 is given by Eq. (4.32). Eliminating V_{LP} and V_{BP} yields $V_N/V_i = H_{0N}H_N$, $V_{BP}/V_i = H_{0BP}H_{BP}$, and $V_{LP}/V_i = H_{0LP}H_{LP}$, where

$$f_z = f_0 = f_1 \qquad Q = R_3/R_2 \tag{4.33a}$$

$$H_{0N} = H_{0LP} = -R_2/R_1 \qquad H_{0BP} = -R_3/R_1 \tag{4.33b}$$

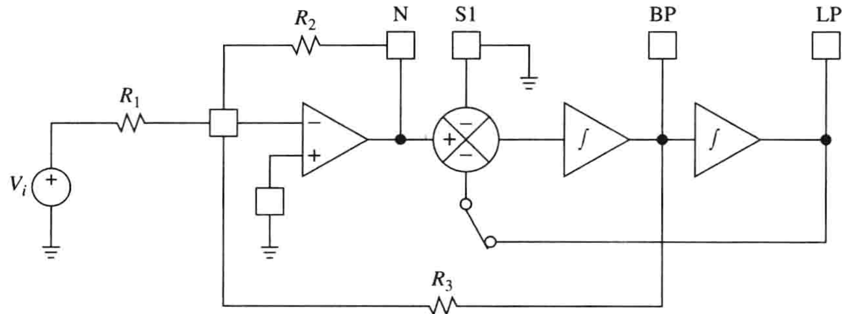


FIGURE 4.36
Basic MF10 connection for the notch, band-pass, and low-pass responses.

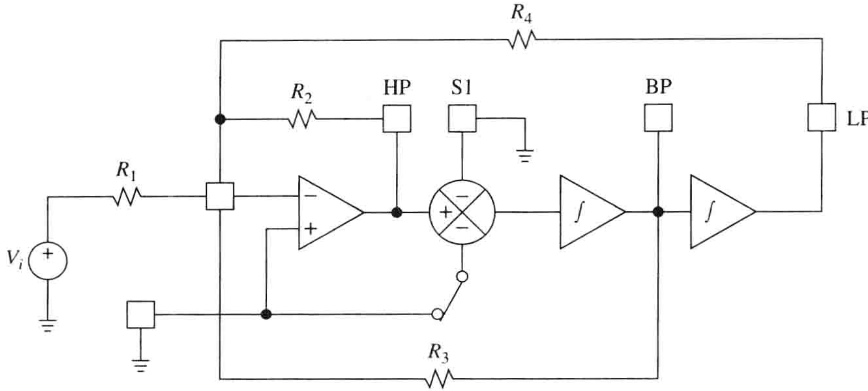


FIGURE 4.37
State-variable configuration using the MF10.

Note that in this mode both f_z and f_0 coincide with the integration unity-gain frequency $f_1 = f_{CK}/100(50)$.

EXAMPLE 4.15. In the circuit of Fig. 4.36, specify suitable resistances for a band-pass response with $f_0 = 1$ kHz, BW $= 50$ Hz, and $H_{0BP} = 20$ dB.

Solution. Impose $R_3/R_2 = Q = f_0/\text{BW} = 10^3/50 = 20$, and $R_3/R_1 = |H_{0BP}| = 10^{20/20} = 10$. Pick $R_1 = 20$ kΩ, $R_2 = 10$ kΩ, $R_3 = 200$ kΩ, $f_{CK} = 100$ kHz, and tie the 50/100/CL pin to ground to make $f_1 = f_{CK}/100$.

The mode of Fig. 4.37 is referred to as the *state-variable* mode because it provides the high-pass, band-pass, and low-pass responses by direct consecutive integrations. One can readily show (see Problem 4.33) that, if $f \ll f_{CK}$, the circuit gives $V_{HP}/V_i = H_{0HP}H_{HP}$, $V_{BP}/V_i = H_{0BP}H_{BP}$, and $V_{LP}/V_i = H_{0LP}H_{LP}$, where

$$f_0 = f_1\sqrt{R_2/R_4} \qquad Q = (R_3/R_2)\sqrt{R_2/R_4} \tag{4.34a}$$

$$H_{0HP} = -R_2/R_1 \qquad H_{0BP} = -R_3/R_1 \qquad H_{0LP} = -R_4/R_1 \tag{4.34b}$$

A distinctive feature of this mode is that f_0 can be tuned independently of the integration unity-gain frequency $f_1 = f_{CK}/100(50)$ by means of the ratio R_2/R_4, a feature we shall exploit in cascade design. Since the summing amplifier is now inside the integrator loop, the frequency limitations of its open-loop gain are likely to cause Q enhancement, a subject that will be addressed in Chapter 6. Suffice it to say here that this enhancement can be compensated by placing a phase-lead capacitance on the order of 10 pF to 100 pF in parallel with R_4.

By combining the high-pass and low-pass responses with an external summing amplifier, in the familiar manner of Fig. 4.38, the notch response is synthesized. One can readily show (see Problem 4.33) that, if $f \ll f_{CK}$, the circuit gives

$$\frac{V_o}{V_i} = H_{0N}\frac{1-(f/f_z)^2}{1-(f/f_0)^2 + (jf/f_0)/Q}$$

$$f_0 = f_1\sqrt{R_2/R_4} \qquad f_z = f_1\sqrt{R_H/R_L} \qquad Q = R_3/R_2\sqrt{R_2/R_4} \tag{4.35a}$$

$$H_{0N} = \frac{R_G R_4}{R_L R_1} \qquad H_{0HP} = -\frac{R_2}{R_1} \qquad H_{0BP} = -\frac{R_3}{R_1} \qquad H_{0LP} = -\frac{R_4}{R_1} \tag{4.35b}$$

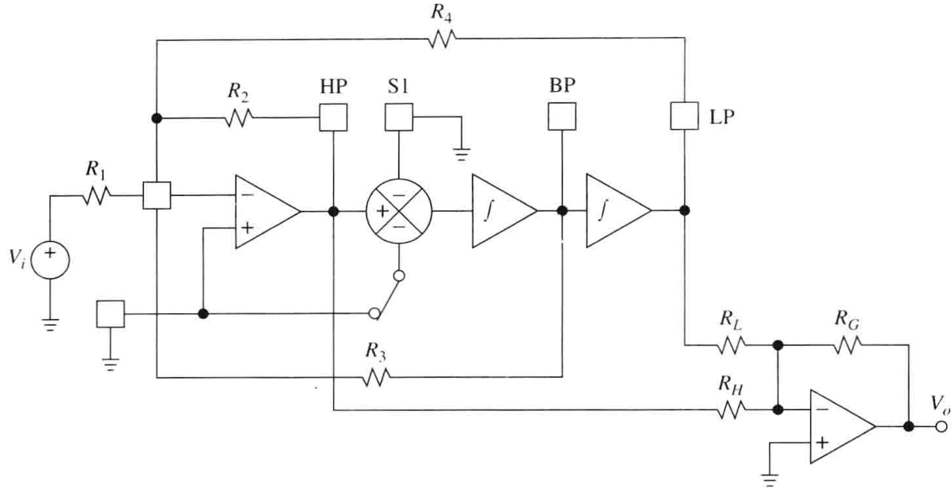


FIGURE 4.38
MF10 with an external op amp to provide the notch response.

Depending on how one specifies the various resistances, the notch can be of the high-pass or low-pass type, and it can be utilized in the synthesis of Cauer filters. When cascading, the high-pass and low-pass outputs of a given section can be combined by means of the input amplifier of the following section, thus reducing the number of external op amps to one, that of the last section.

Cascade Design

Dual-integrator-loop sections can be cascaded to synthesize higher-order filters. If we drive all sections with the same clock, the overall filter will be programmable, since varying f_{CK} will translate all responses up or down the frequency spectrum without affecting their Qs or gains. The resonance frequencies of the individual sections may require shifting with respect to the characteristic frequency of the overall filter. This is accomplished by means of R_4, as demonstrated by Eqs. (4.34*a*) and (4.35*a*). Following are a few cascade-design examples; others can be found in the manufacturer's literature.[4]

EXAMPLE 4.16. Using the MF10 filter, design a fourth-order 1.0-dB Chebyshev low-pass filter with $f_c = 2$ kHz and 0-dB dc gain.

Solution. Let $f_{CK} = 100 f_c = 200$ kHz. From Table 4.1 we find that the following individual-stage parameters are needed: $f_{01} = 0.993 f_c$, $Q_1 = 3.559$, $f_{02} = 0.529 f_c$, and $Q_2 = 0.785$. Let section A be the low-Q stage, and section B the high-Q stage, and let us cascade them in this order to maximize filter dynamics. Since both sections require frequency shifting with respect to f_c, we use the configuration of Fig. 4.37.

By Eq. (4.34), $\sqrt{R_{2A}/R_{4A}} = 0.529$, or $R_{2A}/R_{4A} = 0.2798$; $R_{3A}/R_{2A} = Q_A/\sqrt{R_{2A}/R_{4A}} = 0.785/0.529 = 1.484$; $R_{4A}/R_{1A} = |H_{0LPA}| = 1$. Let $R_{1A} = R_{4A} = 20\,\text{k}\Omega$. Then, $R_{2A} = 5.60$ kΩ and $R_{3A} = 8.30$ kΩ. Likewise, we find $R_{1B} = R_{4B} = 20$ kΩ,

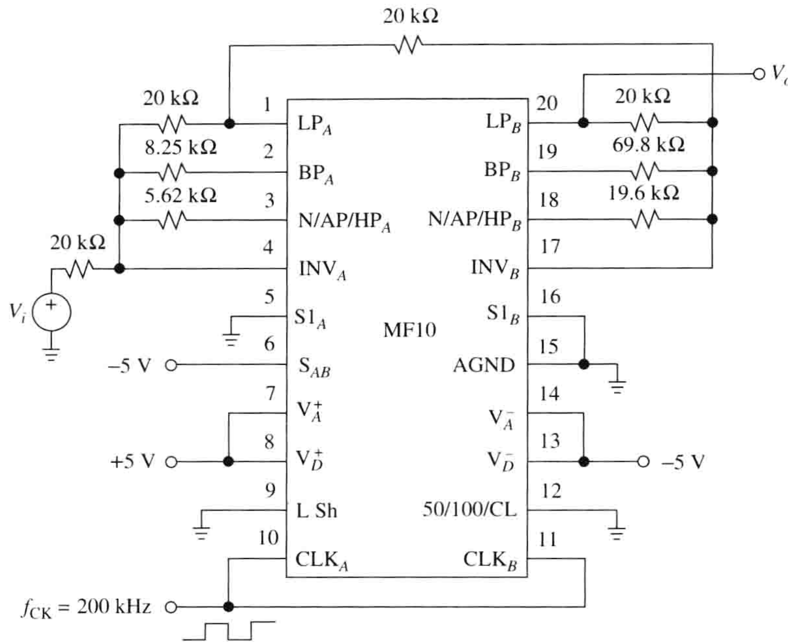


FIGURE 4.39
Fourth-order, 1-dB, 2-kHz Chebyshev low-pass filter of Example 4.16.

$R_{2B} = 19.7$ kΩ, and $R_{3B} = 70.7$ kΩ. The final circuit is shown in Fig. 4.39, where the resistances have been rounded off to 1% standard values. For optimum performance, bypass the power supplies with 0.1-μF disk capacitors right at the supply pins.

EXAMPLE 4.17. Design an elliptic low-pass filter meeting the following specifications: $f_c = 1$ kHz, $f_s = 2$ kHz, $A_{\max} = 1.0$ dB, $A_{\min} = 50$ dB, and 0-dB dc gain.

Solution. The aforementioned FILDES program indicates that we need a fourth-order filter with the following individual-stage parameters:

$$f_{01} = 0.5650 \text{ kHz} \qquad f_{z1} = 2.1432 \text{ kHz} \qquad Q_1 = 0.8042$$

$$f_{02} = 0.9966 \text{ kHz} \qquad f_{z2} = 4.9221 \text{ kHz} \qquad Q_2 = 4.1020$$

Moreover, the actual attenuation at 2 kHz is 51.9 dB.

Use the notch configuration of Fig. 4.38 with $f_{CK} = 100 f_c = 100$ kHz. Design section A first. Let $R_{1A} = 20$ kΩ. Imposing $|H_{0LPA}| = 1$ V/V gives $R_{4A} = R_{1A} = 20$ kΩ. To achieve the given f_{01}, we need $R_{2A}/R_{4A} = 0.5650^2$, or $R_{2A} = 6.384$ kΩ. To achieve the given Q_1, we need $R_{3A} = R_{2A} Q_1/\sqrt{R_{2A}/R_{4A}} = 6.384 \times 0.8042/0.5650 = 9.087$ kΩ. Let $R_{LA} = 20$ kΩ, so that to achieve the given f_{z1} we need $R_{HA}/R_{LA} = 2.1432^2$, or $R_{HA} = 91.87$ kΩ.

Now design section B using the input amplifier of section B to combine the high-pass and low-pass responses of section A. Imposing $|H_{0LPB}| = 1$ V/V gives $R_{4B} = R_{LA} =$ 20 kΩ. Repeating similar calculations, we obtain $R_{2B} = 19.86$ kΩ, $R_{3B} = 81.76$ kΩ, $R_{LB} = 20$ kΩ, and $R_{HB} = 484.5$ kΩ. The last notch requires an external op amp with $R_G = R_{LB} = 20$ kΩ to ensure a 0-dB dc gain. The final circuit is shown in Fig. 4.40, where the resistances have been rounded off to 1% standard values.

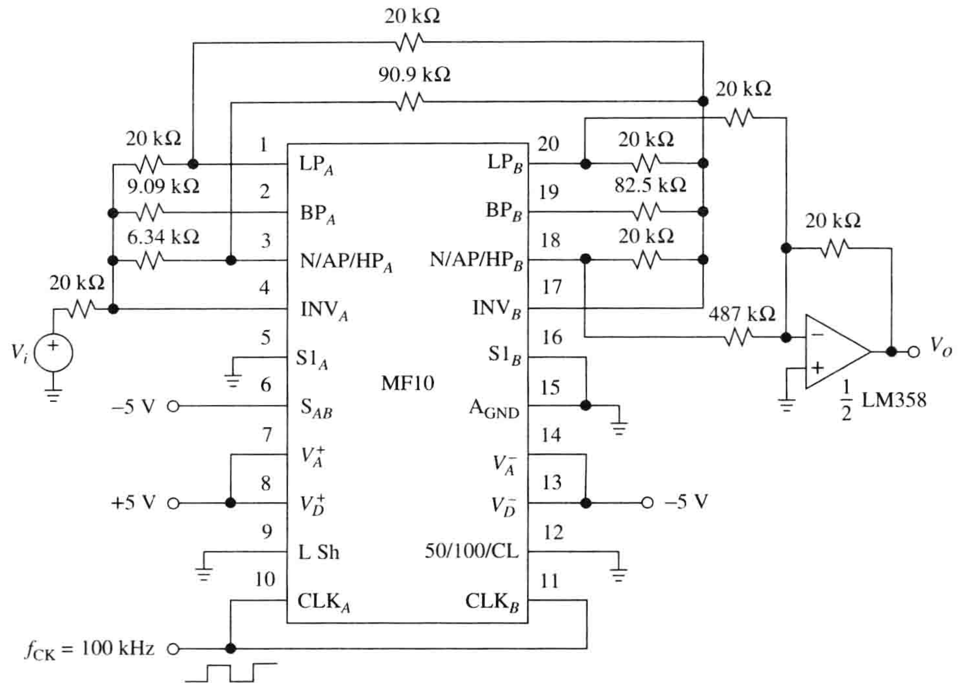


FIGURE 4.40
Fourth-order, 1-dB, 1-kHz elliptic low-pass filter of Example 4.17.

PROBLEMS

4.1 Filter approximations

4.1 (*a*) Find n for a low-pass Butterworth filter with $A_{\max} = 1$ dB, $A_{\min} = 20$ dB, and $\omega_s/\omega_c = 1.2$. (*b*) Find the actual value of $A(\omega_s)$. (*c*) Find $A_{\max}$ so that $A(\omega_s) = 20$ dB exactly.

4.2 Using Eq. (4.5), find n for a low-pass Chebyshev response with the same specifications as the Butterworth response of Example 4.1.

4.3 Using Eq. (4.6), find the passband frequencies at which the gain of a seventh-order 0.5-dB Chebyshev filter exhibits its peaks and valleys, as well as the gain at $2\omega_c$, $10\omega_c$.

4.4 (*a*) Sketch the magnitude plots of the Butterworth and Chebyshev responses for $n = 5$ and $A_{\max} = 1$ dB. (*b*) Compare the attenuations provided at $\omega = 2\omega_c$.

4.5 The normalized third-order Butterworth low-pass response is $H(s) = 1/(s^3 + 2s^2 + 2s + 1)$. (*a*) Verify that it satisfies Eq. (4.4) with $\epsilon = 1$. (*b*) Show that if $k_1 = 0.14537$ and $k_2 = 2.5468$, the single-op-amp filter of Fig. P4.5 implements the third-order Butterworth response with $\omega_c = 1/RC(k_1k_2)^{1/3}$. (*c*) Specify components for $f_c = 1$ kHz.

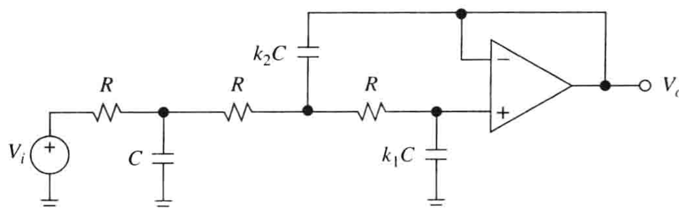


FIGURE P4.5

4.2 Cascade design

4.6 The normalized fourth-order Butterworth low-pass response can be factored as $H(s) = [s^2 + s(2 - 2^{1/2})^{1/2} + 1]^{-1} \times [s^2 + s(2 + 2^{1/2})^{1/2} + 1]^{-1}$. (*a*) Verify that it satisfies the condition of Eq. (4.4) with $\epsilon = 1$. (*b*) Design a fourth-order Butterworth low-pass filter with $f_c = 880$ Hz and $H_0 = 0$ dB.

4.7 A drawback of the implementation of Fig. 4.7 is its high capacitance spread, especially in the high-Q stage. This can be avoided by using $K > 1$. Redesign the filter so that the capacitance spread is kept below 10 while still ensuring 0-dB dc gain.

4.8 The smoothing filter of Fig. 4.7 is adequate for moderate performance requirements. Ultra-high fidelity audio applications require lower passband ripple and even sharper cutoff characteristics. For a 40-kHz sampling rate, these demands can be met[9] with a tenth-order 0.25-dB Chebyshev low-pass filter having $f_c = 15$ kHz. Such a filter provides $A(20 \text{ kHz}) = 50.5$ dB with a -3-dB frequency of 15.35 kHz. The individual-stage parameters are: $f_{01} = 3.972$ kHz, $Q_1 = 0.627$, $f_{02} = 7.526$ kHz, $Q_2 = 1.318$, $f_{03} = 11.080$ Hz, $Q_3 = 2.444$, $f_{04} = 13.744$ kHz, $Q_4 = 4.723$, $f_{05} = 15.158$ kHz, and $Q_5 = 15.120$. Design such a filter and show your final circuit.

4.9 Using equal-component *KRC* sections, design a fifth-order Bessel low-pass filter with $f_c = 1$ kHz and $H_0 = 0$ dB.

4.10 Using *KRC* sections with $C_1 = C_2$ and $R_A = R_B$, design a seventh-order Butterworth low-pass filter with $f_c = 1$ kHz and $H_0 = 20$ dB.

4.11 Design a fifth-order 1.0-dB Chebyshev high-pass filter with $f_c = 360$ Hz and high-frequency gain H_0 adjustable from 0 to 20 dB. Use equal capacitances throughout.

4.12 A band-pass filter is to be designed with center frequency $f_0 = 300$ Hz, $A(300 \pm 10 \text{ Hz}) = 3$ dB, $A(300 \pm 40 \text{ Hz}) \geq 25$ dB, and resonance gain $H_0 = 12$ dB. These specifications[3] can be met with a sixth-order cascade filter having the following individual-stage parameters: $f_{01} = 288.0$ Hz, $Q_1 = 15.60$, $H_{0BP1} = 2.567$ V/V; $f_{02} = 312.5$ Hz, $Q_2 = 15.60$, $H_{0BP2} = 2.567$ V/V; $f_{03} = 300.0$ Hz, $Q_3 = 15.34$, $H_{0BP3} = 1.585$ V/V. Design such a filter using three individually tunable multiple-feedback stages.

4.13 Complete the design of Example 4.7, and show the final circuit.

4.14 Using the cascade-design approach, along with the FILDES program, design a 0.5-dB Chebyshev low-pass filter with a cutoff frequency of 10 kHz, a stopband frequency of 20 kHz, a minimum stopband attenuation of 60 dB, and a dc gain of 12 dB. Then, run a PSpice simulation of your circuit, showing the magnitude Bode plots of the individual-stage responses as well as the overall response.

4.15 Recalculate the component values of the sixth-order band-pass filter of Fig. 4.11 (the values shown have been rounded off to 1% standard values). Then (*a*) use PSpice, along with pseudo-ideal op amps, to plot the ac response of the circuit, measure the filter parameters with the cursor, and compare with the intended values. (*b*) Repeat, but using 741 op amps. This will expose you to the effect of limited op amp dynamics upon filters, a subject to be addressed in Chapter 6.

4.3 Generalized impedance converters

4.16 In both circuits of Fig. 4.15 let all resistances be 1 kΩ and all capacitances be 1 nF. Use PSpice to display the Bode plots of both impedances (call them Z_L and Z_D). Compare slopes and phases, and justify. Predict the frequency at which $|Z_L| = |Z_D|$, measure it with the cursor, and compare.

4.17 (*a*) Using the DABP filter of Fig. 4.17*b*, along with a summing amplifier, design a second-order notch filter with $f_z = 120$ Hz and $Q = 20$. (*b*) Suitably modify the circuit of part (*a*) for a second-order all-pass filter with a gain of 20 dB.

4.18 It is desired to design a band-pass filter with $f_0 = 1$ kHz, $A(f_0 \pm 10 \text{ Hz}) = 3$ dB, and $A(f_0 \pm 40 \text{ Hz}) \geq 20$ dB. Such a filter[3] can be implemented by cascading two second-order band-pass stages with $f_{01} = 993.0$ Hz, $f_{02} = 1007$ Hz, and $Q_1 = Q_2 = 70.7$. Design an implementation using the DABP filter of Fig. 4.17*b*. Make provision for frequency tuning of the individual stages.

4.19 (*a*) Show that Eq. (4.12) holds also for the *D* element of Fig. 4.16. (*b*) Using this element, along with the *RLC* prototype of Fig. 4.18*a*, design a low-pass filter with $f_0 = 800$ Hz and $Q = 4$.

4.20 Provided $R = \sqrt{2L/C}$, the circuit of Fig. P4.20 yields a third-order, high-pass Butterworth response with −3-dB frequency $\omega_c = 1/\sqrt{2LC}$. (*a*) Specify suitable components for $f_c = 1$ kHz. (*b*) Convert the circuit to a GIC realization.

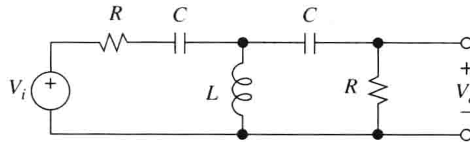


FIGURE P4.20

4.21 Show that the circuit of Fig. P4.21 simulates a grounded inductance $L = R_1R_3R_4C/R_2$.

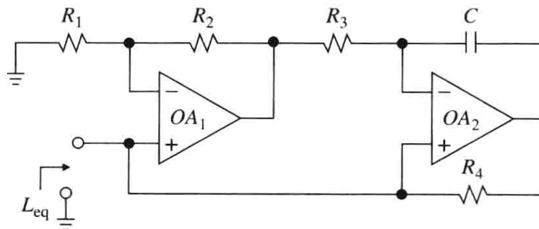


FIGURE P4.21

4.22 The circuit of Fig. P4.22 simulates an impedance Z_1 proportional to the reciprocal of Z_2. Called a *gyrator*, it finds application as an inductance by letting Z_2 be a capacitance. (*a*) Show that $Z_1 = R^2/Z_2$. (*b*) Using this circuit, design a second-order band-pass filter with $f_0 = 1$ Hz, $Q = 10$, and zero output impedance. What is the resonance gain of your circuit?

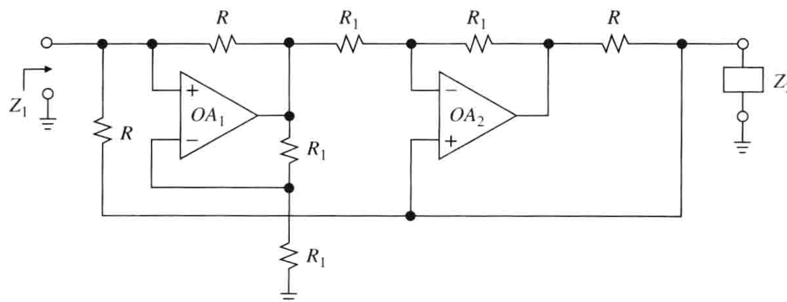


FIGURE P4.22

4.23 Let all resistances in the circuit of Fig. P4.22 be 1 kΩ. (*a*) Use PSpice, along with pseudo-ideal op amps, to plot $|Z_1|$ if Z_2 is a 1-nF capacitance. (*b*) Repeat, if Z_2 is a 100-μH inductance. Compare and comment. (*c*) Repeat, but using 741 op amps. Based on what you see, specify the range of frequencies over which each of the 741-based circuits performs reasonably close to the ideal.

4.4 Direct design

4.24 It is desired to design a seventh-order 0.5-dB Chebyshev low-pass filter with a −3-dB frequency of 10 kHz. From Table 4.2 we find the *RLC* element values shown in Fig. P4.24. Using this ladder as a prototype, design an FDNR implementation.

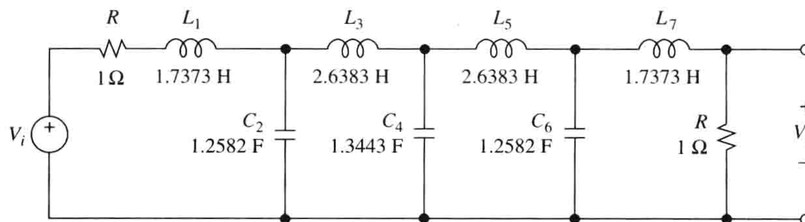


FIGURE P4.24

4.25 Using GICs and the information of Table 4.2, design a seventh-order 1-dB Chebyshev high-pass filter with $f_c = 500$ Hz.

4.26 Use PSpice, along with pseudo-ideal op amps, to plot the ac response of the seventh-order low-pass filter of Fig. 4.21. Measure the filter parameters with the cursor, and compare with the intended values.

4.5 The switched capacitor

4.27 Find a relationship beween V_o and V_1 and V_2 in the circuits of Fig. P4.27 for $f \ll f_{CK}$, and give the circuits' descriptive names.

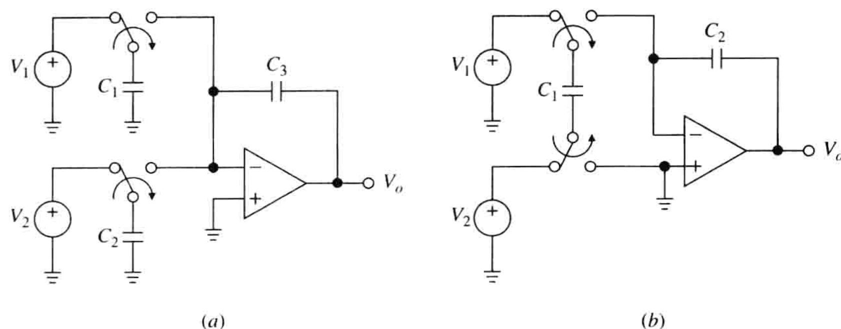


FIGURE P4.27

4.28 Find the transfer function of the circuits of Fig. P4.28 for $f \ll f_{CK}$, and give the circuits' descriptive names.

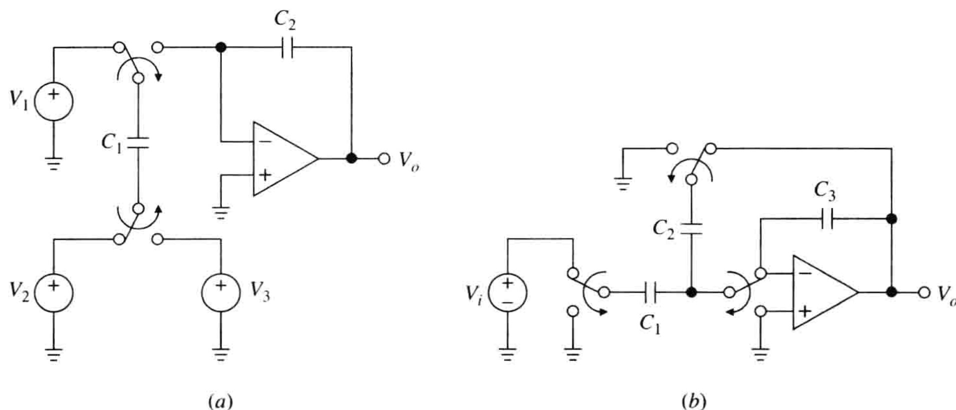


FIGURE P4.28

4.29 (a) Assuming $f \ll f_{CK}$, show that the circuit of Fig. P4.29 gives the notch response. (b) Assuming $f_{CK} = 100$ kHz, specify capacitances for a 1-kHz notch with $Q = 10$.

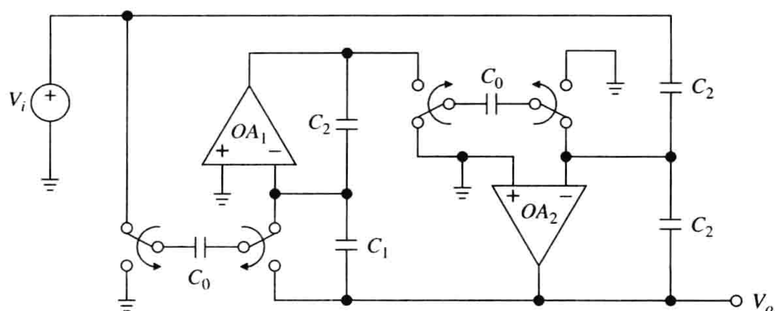


FIGURE P4.29

4.30 (*a*) Assuming $f_{\rm CK} = 250$ kHz in the circuit of Fig. 4.29, specify capacitances for a band-pass response with $f_0 = 2$ kHz and BW $= 1$ kHz. (*b*) Repeat, but for BW $= 100$ Hz. What do you conclude about the capacitance spread as a function of Q?

4.31 (*a*) Derive Eq. (4.25). (*b*) Assuming $f_{\rm CK} = 200$ kHz, specify capacitances in the circuit of Fig. 4.30 for $f_0 = 1$ kHz and $Q = 10$. (*c*) Repeat, but for $Q = 100$. Comment on the capacitance spread.

4.32 Using Table 4.2, but with $C_1, L_2, C_3, \ldots$ as column headings, design a fifth-order 0.1-dB Chebyshev low-pass SC ladder filter with $f_c = 3.4$ kHz and $f_{\rm CK} = 128$ kHz.

4.7 Universal SC filters

4.33 Derive Eqs. (4.34) and (4.35).

4.34 Consider the circuit obtained from that of Fig. 4.36 by removing R_1, lifting the S1 pin off ground, and applying V_i to S1, so that only two resistances are used. (*a*) Sketch the modified circuit and show that $V_{\rm BP}/V_i = -QH_{\rm BP}$ and $V_{\rm LP}/V_i = -H_{\rm LP}$, with f_0 and Q given by Eq. (4.33*a*). (*b*) Specify resistances for $f_0 = 500$ Hz and $Q = 10$.

4.35 The MF10 configuration of Fig. P4.35 provides the notch, band-pass, and low-pass responses, with the notch frequency f_z and the resonance frequency f_0 independently tunable by means of the resistance ratio R_2/R_4. Find expressions for f_0, f_z, Q, and the low-frequency gain.

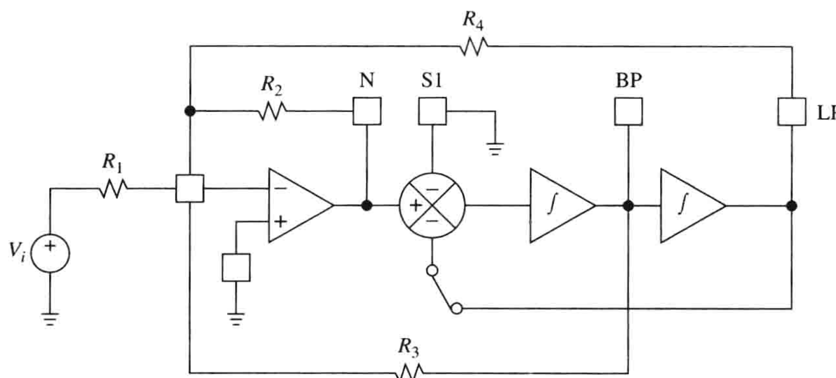


FIGURE P4.35

4.36 If in the circuit of Fig. P4.35 we lift the S1 input off ground and we connect it to V_i, with everything else remaining the same, then the output of the leftmost op amp changes from the notch to the all-pass response, with the numerator and denominator Qs separately adjustable. Assuming $f \ll f_{\rm CK}$, find f_0, the numerator and denominator Qs, and the gain.

4.37 Using the MF10 in the configuration of Problem 4.34, design a minimum-component fourth-order Butterworth low-pass filter with $f_c = 1$ kHz and 20-dB dc gain.

4.38 A fourth-order 0.5-dB Chebyshev band-pass filter is to be designed with $f_0 = 2$ kHz and BW $= 1$ kHz. Using the FILDES program, it is found that the cascade realization

requires the following individual-stage parameters: $f_{01} = 1554.2$ Hz, $f_{02} = 2473.6$ Hz, and $Q_1 = Q_2 = 2.8955$. Design such a filter using the MF10.

4.39 A fourth-order 1.0-dB Chebyshev notch filter with $f_0 = 1$ kHz is to be implemented by cascading two second-order sections with $f_{01} = 1.0414 f_0$, $f_{02} = 0.9602 f_0$, $f_{z1} = f_{z2} = f_0$, and $Q_1 = Q_2 = 20.1$. Design such a filter using the MF10.

4.40 It is desired to design a 0.5-dB elliptic band-pass filter with a center frequency $f_0 = 2$ kHz, a passband of 100 Hz, a stopband of 300 Hz, and a minimum stopband attenuation of 20 dB. Using the FILDES program, it is found that this filter requires a fourth-order implementation with the following individual-stage parameters: $f_{01} = 1.948$ kHz, $f_{z1} = 1.802$ kHz, $f_{02} = 2.053$ kHz, $f_{z2} = 2.220$ kHz, and $Q_1 = Q_2 = 29.48$. Moreover, the actual attenuation at the stopband edges is 21.5 dB. Design such a filter using the MF10 and an external op amp.

4.41 Using two MF10s, design an eighth-order 0.1-dB Chebyshev high-pass filter with $f_c = 500$ Hz and 0-dB high-frequency gain.

REFERENCES

1. L. P. Huelsman, *Active and Passive Analog Filter Design: An Introduction,* McGraw-Hill, New York, 1993.
2. K. Lacanette, "A Basic Introduction to Filters: Active, Passive, and Switched-Capacitor," Application Note AN-779, *Linear Applications Handbook,* National Semiconductor, Santa Clara, CA, 1994.
3. A. B. Williams and F. J. Taylor, *Electronic Filter Design Handbook: LC, Active, and Digital Filters,* 2d ed., McGraw-Hill, New York, 1988.
4. K. Lacanette, ed., *Switched Capacitor Filter Handbook,* National Semiconductor, Santa Clara, CA, 1985.
5. M. Steffes, "Advanced Considerations for Gain and Q Sequencing in Multistage Low-pass Active Filters," *EDN,* Oct. 4, 2010 (downloadable from www.edn.com).
6. A. S. Sedra and J. L. Espinoza, "Sensitivity and Frequency Limitations of Biquadratic Active Filters," *IEEE Trans. Circuits and Systems,* Vol. CAS-22, No. 2, Feb. 1975.
7. L. T. Bruton and D. Treleaven, "Active Filter Design Using Generalized Impedance Converters," *EDN,* Feb. 5, 1973, pp. 68–75.
8. L. Weinberg, *Network Analysis and Synthesis,* McGraw-Hill, New York, 1962.
9. D. J. M. Baezlopez, *Sensitivity and Synthesis of Elliptic Functions,* Ph.D. Dissertation, University of Arizona, 1978.
10. H. Chamberlin, *Musical Applications of Microprocessors,* 2d ed., Hayden Book Company, Hasbrouck Heights, NJ, 1985.
11. A. B. Grebene, *Bipolar and MOS Analog Integrated Circuit Design,* John Wiley & Sons, New York, 1984.
12. P. E. Allen and E. Sanchez-Sinencio, *Switched Capacitor Circuits,* Van Nostrand Reinhold, New York, 1984.
13. R. Gregorian and G. C. Temes, *Analog MOS Integrated Circuits for Signal Processing,* John Wiley & Sons, New York, 1986.

5

STATIC OP AMP LIMITATIONS

如果你有机会用目前讲过的运算放大器电路做实验，可能会注意，只要运算放大器工作在恰当的频率和适度的闭环增益下，那么其实际工作特性和基于理想运算放大器模型预测的特性往往就是相当一致的。然而，随着频率或增益的增大，电路的频率响应和瞬态响应会进一步恶化。运算放大器在频域和时域的工作特性，统称为运放的动态特性，将在第 6 章研究。

即使保持工作频率相当低，另一些限制也会影响运算放大器。这些限制统称为输入参考误差，在高直流增益应用场合尤其值得注意。其中最常见的几种是输入偏置电流 I_B、输入失调电流 I_{OS}、输入失调电压 V_{OS}，以及交流噪声密度 e_n 和 i_n。相关的研究论题是热漂移、共模抑制比 CMRR、电源抑制比 PSRR 和增益非线性

度。这些非理想特性一般无法通过负反馈得到改善，需要通过其他方法在一对一的基础上减小它们的影响。为了使运算放大器工作正常，还必须遵守一定的工作条件限制，其中包括最大工作温度、最大供电电压、最大功耗、输入共模电压范围和输出短路电流。本章介绍除交流噪声外的所有这些限制。交流噪声将在第 7 章讨论。

虽然所有这些可能听起来让人感到沮丧，但决不应该放弃对理想运算放大器模型的信心，因为它对于初步理解大多数电路仍是一个有力的工具。只有在第二个分析过程中，用户才通过更加细致的分析来研究实际限制因素的影响，从而鉴别问题的来源，并根据需要采取修正措施。

为了方便研究，应该在集中讨论一种限制的同时，假设运算放大器的其他方面都是理想的。实际上，所有限制都是同时存在的。然而，单独地评估它们的影响有助于我们更好地衡量它们的相对重要性，并且找出对当前应用场合最关键的几种限制。

原则上，只要知道运算放大器的内部电路图和工艺参数，每种限制都可以通过手算或者计算机仿真得到预计。另一种方法是把这个器件当作一个黑匣子，根据数据手册上的有用信息对它建立模型，并预测其工作特性。如果实际性能不符合设计目标，设计者要么改变电路，要么选择不同的器件，或者两种方法同时采用，直到得到一个满意的结果为止。

要成功地应用模拟电路，关键是正确地理解数据手册的信息。这个过程会通过使用附录 5A 的 741 数据手册来说明。限于篇幅，无法在此介绍其他器件的数据手册。幸运的是，现在每个人只要在网站上使用器件的部分名字（如“741”“OP77”等）作为关键词搜索，基本上就能找到任何数据手册。

本章重点

本章开始先简要介绍几种代表性的运算放大器工艺（比如三极管、JFET 和 CMOS）的内部电路图。为了有效地选择和利用器件，用户需要对内部工作如何影响实际器件的各种限制有一个基本了解。

接着，本章介绍输入偏置电流以及它们在电路中引入的误差。为帮助用户选择器件，本章讨论常见的拓扑和工艺技巧，用于减小输入偏置电流。

本章随后介绍输入失调电压，这是一个相当复杂的参数，但它简化了许多非理想特性的建模，比如内部元件失调、热漂移、对电源和共模输入电压变化的敏感性，以及有限增益。为帮助用户选择器件，本章讨论常见的拓扑和工艺技巧，用于减小输失调电压。

输入偏置电流和输入失调电压共同产生所有的输入误差，因此接下来的任务就是阐述消除这些误差的常见技巧。

本章结尾讲述最大额定值、输出短路保护以及输入电压范围、输出电压摆幅和轨对轨性能的概念，这些概念在当今的低电压供电系统中十分重要。

5.1 SIMPLIFIED OP AMP CIRCUIT DIAGRAMS

Even though the data sheets provide all the information the user needs to know, a basic familiarity with op amp technologies/topologies will help select the optimal device for a given application. Based on technology, op amps come in three broad categories: (*a*) *bipolar* op amps, (*b*) *JFET-input* op amps, also called *biFET* op amps, and (*c*) *CMOS* op amps (some products combine the best of both bipolar and CMOS on the same chip and are thus called *biCMOS* op amps). Based on topology, two main categories are prevalent: (*a*) *voltage-feedback* op amps (VFAs), by far the most popular, and *current-feedback* op amps (CFAs), a more recent and generally faster category that will be covered in Chapter 6. Other, more specialized topologies are available, such as Norton amplifiers and operational transconductance amplifiers (OTAs), but will not be addressed here.

We begin with the simplified bipolar diagram of Fig. 5.1, inspired by the industry-standard 741 and encountered, in variant forms, in a multitude of other VFAs (in this connection you are urged to search the web for the classic tutorial study by J. E. Solomon,[1] probably the most widely read article on the subject). The diagram shows three basic building blocks, namely, the first or *input stage*, the second or *intermediate stage*, and the third or *output stage*.

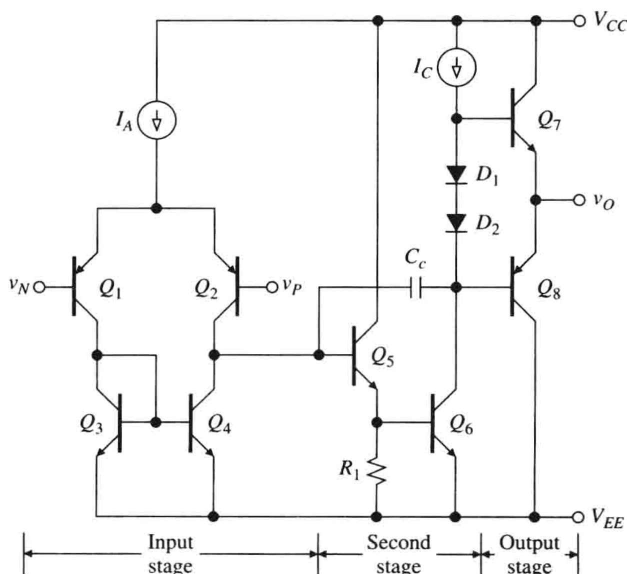


FIGURE 5.1
Simplified circuit diagram of a typical bipolar op amp.

The Input Stage

This stage senses any imbalance between the inverting and noninverting input voltages v_N and v_P, and converts it to a single-ended output current i_{O1} according to

$$i_{O1} = g_{m1}(v_P - v_N) \tag{5.1}$$

where g_{m1} is the input-stage *transconductance*. This stage is designed to also provide high input impedance and draw negligible input currents. As shown again in Fig. 5.2*a*, the input stage consists of two matched transistor pairs, namely, the *differential pair* Q_1 and Q_2, and the *current mirror* Q_3 and Q_4.

The input-stage bias current I_A splits between Q_1 and Q_2. Ignoring transistor base currents and applying KCL, we have

$$i_{C1} + i_{C2} = I_A \tag{5.2}$$

For a *pnp* transistor, the collector current i_C is related to its emitter-base voltage drop v_{EB} by the well-known exponential law,

$$i_C = I_s \exp(v_{EB}/V_T) \tag{5.3}$$

where I_s is the *collector saturation current* and V_T the *thermal voltage* ($V_T \cong 26$ mV at room temperature). Assuming matched BJTs ($I_{s1} = I_{s2}$), we can write

$$\frac{i_{C1}}{i_{C2}} = \exp\left(\frac{v_{EB1} - v_{EB2}}{V_T}\right) = \exp\left(\frac{v_P - v_N}{V_T}\right) \tag{5.4}$$

where we have used $v_{EB1} - v_{EB2} = v_{E1} - v_{B1} - (v_{E2} - v_{B2}) = v_{B2} - v_{B1} = v_P - v_N$.

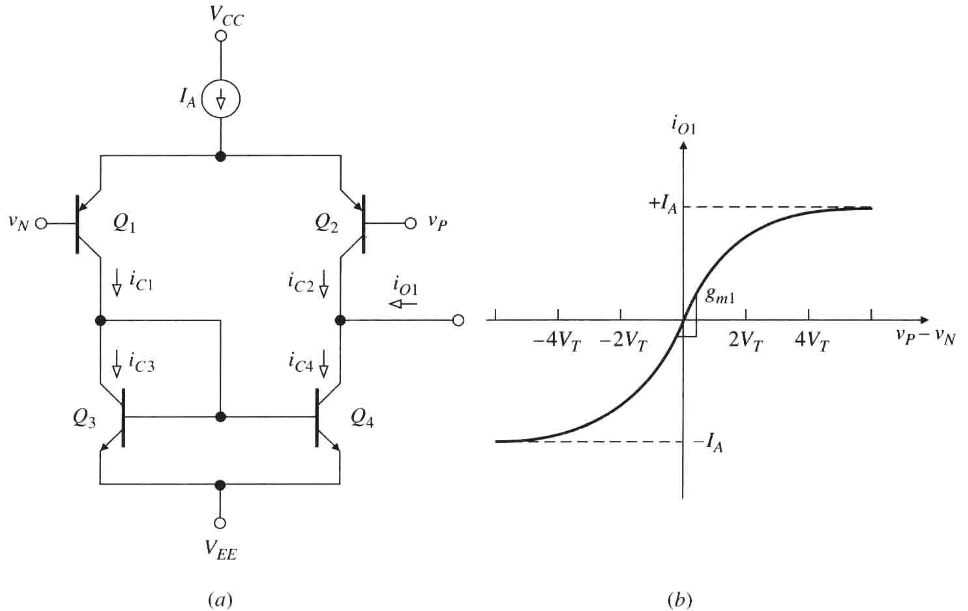


FIGURE 5.2
Input stage and its transfer characteristic.

In response to i_{C1}, Q_3 develops a certain base-emitter voltage drop v_{BE3}. Since $v_{BE4} = v_{BE3}$, Q_4 is forced to draw the same current as Q_3, or $i_{C4} = i_{C3}$; hence the designation current mirror. But, $i_{C3} = i_{C1}$, so the first-stage output current is, by KCL, $i_{O1} = i_{C4} - i_{C2} = i_{C1} - i_{C2}$. Solving Eqs. (5.2) and (5.4) for i_{C1} and i_{C2}, and then taking their difference, we get

$$i_{O1} = I_A \tanh \frac{v_P - v_N}{2V_T} \tag{5.5}$$

This function is plotted in Fig. 5.2*b*.

We observe that under the balanced condition $v_P = v_N$, I_A splits equally between Q_1 and Q_2, thus yielding $i_{O1} = 0$. However, any imbalance between v_P and v_N will divert more of I_A through Q_1 and less through Q_2, or vice versa, thus yielding $i_{O1} \neq 0$. For sufficiently small imbalances, also referred to as *small-signal conditions,* the transfer characteristic is approximately linear and is expressed by Eq. (5.1). The slope, or transconductance, is found as $g_{m1} = di_{O1}/d(v_P - v_N)|_{v_P = v_N}$. The result is

$$g_{m1} = \frac{I_A}{2V_T} \tag{5.6}$$

Overdriving the input stage will eventually force all of I_A through Q_1 and none through Q_2, or vice versa, thus causing i_{O1} to saturate at $\pm I_A$. The overdrive conditions are referred to as *large-signal conditions*. From the figure we see that the onset of saturation occurs for $v_P - v_N \cong \pm 4V_T \cong \pm 100$ mV. As we know, an op amp with negative feedback normally forces v_N to closely track v_P, indicating small-signal operation.

The Second Stage

This stage is made up of the Darlington pair Q_5 and Q_6, and the frequency-compensation capacitance C_c. The Darlington pair is designed to provide additional gain as well as a wider signal swing. The capacitance is designed to stabilize the op amp against unwanted oscillations in negative-feedback applications, a subject to be addressed in Chapter 8. Since C_c is fabricated on-chip, the op amp is said to be *internally compensated.* By contrast, uncompensated op amps require that the compensation network be supplied externally by the user. The 741 op amp is internally compensated. A popular uncompensated contemporary is the 301 op amp.

The Output Stage

This stage, based on the emitter followers Q_7 and Q_8, is designed to provide low output impedance. Though its voltage gain is only approximately unity, its current gain is fairly high, indicating that this stage acts as a power booster for the second-stage output.

Transistors Q_7 and Q_8 are referred to as a *push-pull pair* because in the presence of a grounded output load, Q_7 will source (or push) current to the load during positive output voltage swings, whereas Q_8 will sink (or pull) current from the load during negative swings. The function of the diodes D_1 and D_2 is to develop a pair of

pn-junction voltage drops suitable for biasing Q_7 and Q_8 in the forward-active region and thus minimize crossover distortion at the output.

JFET-Input Op Amps

A potential drawback of bipolar-input op amps are the base currents at the v_P and v_N inputs. The op amp forces these currents automatically through the surrounding circuitry, creating voltage drops that may be unacceptable in certain precision applications (more on this in the next section). One way of alleviating this drawback is by implementing the differential input pair with junction field-effect transistors (JFETs), as shown in Fig. 5.3.

The transfer characteristic of the input stage is still qualitatively similar to that of Fig. 5.2*b*, although g_{m1} is typically much lower because the quadratic characteristic of an FET is less steep than the exponential characteristic of an equally biased BJT. However, the currents at the v_P and v_N inputs are now the gate currents of the JFETs, which at room temperature are orders of magnitude lower than the base currents of BJTs. (The gate current of a JFET is the *leakage current* of the reverse-biased *pn* junction formed by the gate and channel regions.)

The present rendition also shows an alternative second stage implementation, which consists of another differential pair (Q_3-Q_4), along with the corresponding current mirror (Q_5-Q_6). Also shown is an alternative output-stage implementation, which uses the complementary Darlington-like pairs Q_7-Q_{10} and Q_8-Q_9 to provide the required push-pull action at the output while presenting high input resistance so as to limit loading of the second stage. The op amp is frequency compensated by means of C_c.

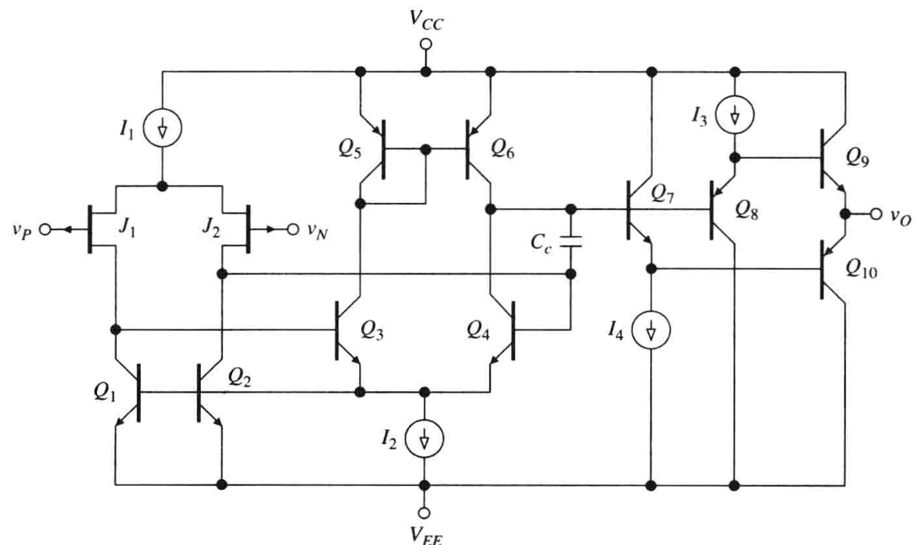


FIGURE 5.3
Simplified circuit diagram of a JFET-input op amp.

The emergence of CMOS technology in digital electronics, in the early to mid-1970s, along with the drive to combine digital as well as analog functions on the same chip, has provided a strong motivation for recasting traditionally bipolar analog functions in forms amenable to CMOS implementation (we have already seen an example in switched-capacitor filters). A transistor's ability to amplify voltages is expressed via a figure of merit known as the *intrinsic voltage gain* $g_m r_o$, where g_m and r_o are, respectively, the transistor's transconductance and output resistance. Compared to BJTs, which are capable of intrinsic gains in the thousands, FETs are notorious for their much lower intrinsic gains, not an encouraging sign if we want to use them to build op amps. However, the MOSFET offers three important advantages: (*a*) it presents virtually infinite input resistance at its gate, essentially eliminating input loading; (*b*) using a technique known as *cascoding*,[3,4] we can raise its effective r_o dramatically to make up for its poor g_m; (*c*) it takes up much less chip area than a BJT, thus allowing for much denser integration. These advantages, combined with clever advances in design and fabrication, have made CMOS op amps competitive with their bipolar counterparts in a number of areas. Shown in Fig. 5.4 are two popular CMOS topologies.

The topology of Fig. 5.4*a* is a CMOS replica of the first two stages of Fig. 5.1. The input stage, consisting of the differential pair M_1-M_2 and current mirror M_3-M_4, provides a transfer characteristic of the type of Fig. 5.2*b*, though typically with a much lower g_{m1}. Moreover, we can exploit the virtually infinite gate resistance of MOSFETs to implement the second stage with a single transistor, M_5. The M_6-M_7-M_8 triplet forms a dual-output current mirror for biasing the M_1-M_2 pair and the M_5 stage. The op amp is frequency compensated by means of the R_c-C_c network.

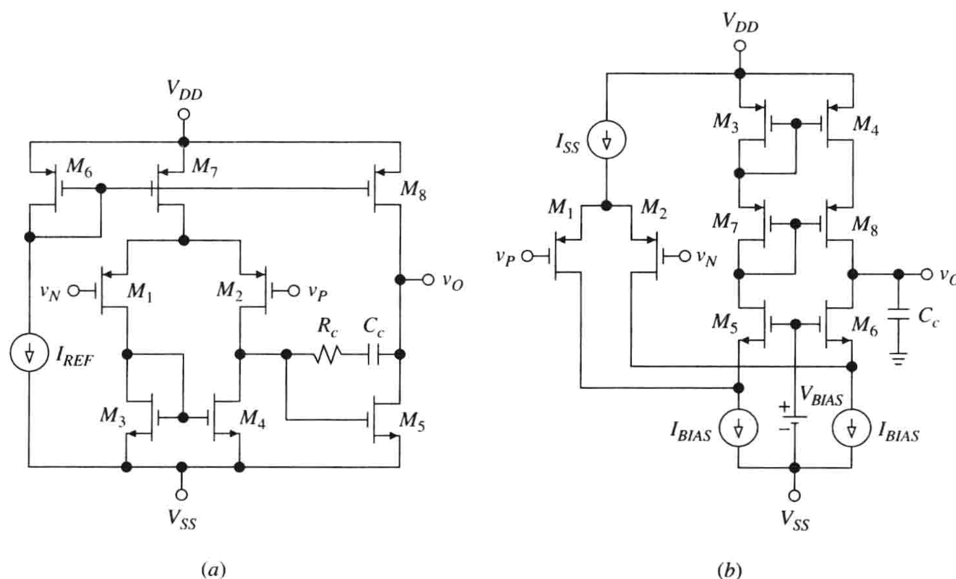


FIGURE 5.4
Two popular CMOS op amp topologies: (*a*) the two-stage and (*b*) the folded-cascode.

Conspicuously absent from the figure is a dedicated output stage. This reflects the fact that nowadays CMOS op amps are most frequently used as subcircuits of larger IC systems, where the output load is known and usually light enough not to warrant a dedicated output stage. (On the contrary, a general-purpose op amp does require a dedicated output stage to confront a wide variety of a priori unknown loads. CMOS op amps with dedicated output stages will be addressed in Section 5.7.)

The topology of Fig. 5.4*b* is said to be of the *single-stage* type because its core is just the differential pair M_1-M_2 and the current mirror M_3-M_4. The remaining FETs serve only the ancillary purpose of raising the intrinsic gain of this basic stage via a technique known as the *cascoding* technique.[3,4] Specifically, M_8 raises r_{o4} and M_6 raises r_{o2} so as to ensure a high overall output resistance R_o and thus maximize the gain, which for this type of op amp is $a = g_{m1} R_o$.

The M_1-M_2 pair senses any imbalance between its gate voltages and converts to an imbalance between its drain currents. The drain-current imbalance is then rerouted upward by the M_5-M_6 pair in *folded* fashion (hence the name), a technique that allows for much wider output voltage swings.[3,4] (In Chapter 8 we shall see that capacitive loads tend to destabilize op amps. Not so in the folded-cascode case, because the load capacitance C_c actually improves stability, a feature that makes folded-cascode op amps particularly suited to switched-capacitor applications.)

SPICE Models

There are various levels at which an op amp can be simulated. In IC design, op amps are simulated at the transistor level,[2] also called the *micromodel* level. Such a simulation requires a detailed knowledge of both the circuit schematic and the fabrication process parameters. However, this proprietary information is not easily accessible to the user. Even so, the level of detail may require excessive computation time or may even cause convergence problems, especially in more complex circuit systems.

To cope with these difficulties, simulations by the user are usually carried out at the *macromodel* level. A macromodel uses a much reduced set of circuit elements to closely match the measured behavior of the finished device while saving considerable simulation time. Like any model, a macromodel comes with limitations, and the user need be aware of the parameters the particular macromodel fails to simulate. Macromodels are available online from virtually every manufacturer. The student version of PSpice used in this book includes a 741 macromodel based on the so-called Boyle model[5] of Fig. 5.5.

At times we may wish to focus on just one particular op amp feature and thus develop an even simpler model on our own. A typical example is offered by the frequency response, to be studied in Chapter 6. Regardless of the model used, a circuit must eventually be breadboarded and tried out in the lab, where its behavior is evaluated in the presence of parasitics and other factors related to actual circuit construction, which computer simulation, unless properly instructed, fails to account for.

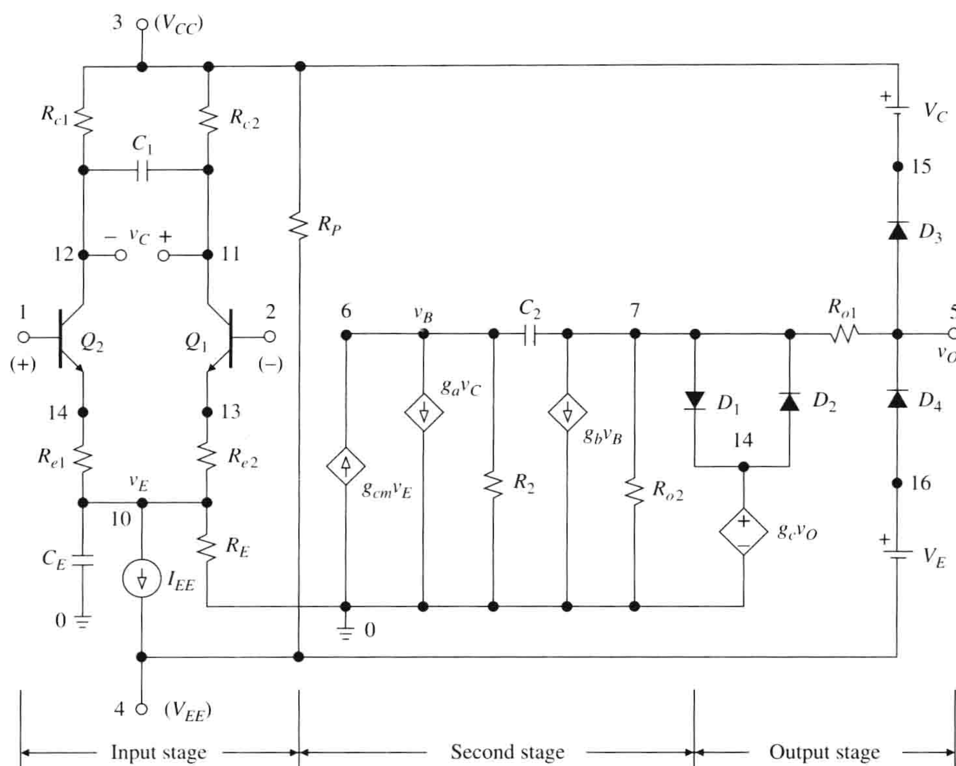


FIGURE 5.5
741 op amp macromodel.

5.2 INPUT BIAS AND OFFSET CURRENTS

Let us now investigate how the input-pin currents affect the performance of an op amp circuit. We shall use the 741 as a vehicle, so we need to take a closer look at the input-stage portion of Fig. 5A.2, which is reproduced in Fig. 5.6 for convenience. The process used to fabricate the 741 optimizes *npn* BJTs at the expense of *pnp* BJTs, which therefore come out with inferior characteristics, such as a much lower current gain β_{Fp}. Had the input stage been fabricated in the simplified form of Fig. 5.1, the *pnp* transistors would have presented intolerably high currents at the v_P and v_N terminals. As seen in Fig. 5.6, the 741 overcomes this limitation by operating the Q_3-Q_4 *pnp* transistors as a common-base pair, and by driving this pair with the high-beta *npn* emitter followers Q_1-Q_2. By this ingenious artifice, the composite structure Q_1-Q_2-Q_3-Q_4 appears as *pnp*-type to the current mirror Q_5-Q_6, but as *npn*-type to nodes v_P and v_N. The addition of the emitter followers halves the transconductance of Eq. (5.6) to $g_{m1} = I_A/(4V_T)$, but it also results in much lower input currents $I_P \cong I_N \cong (I_A/2)/\beta_{Fn}$, owing to the much higher value of β_{Fn}. As we move along, we shall use the following 741 working values:

$$I_A = 19.5\ \mu\text{A} \qquad g_{m1} = 189\ \mu\text{A/V} \tag{5.7}$$

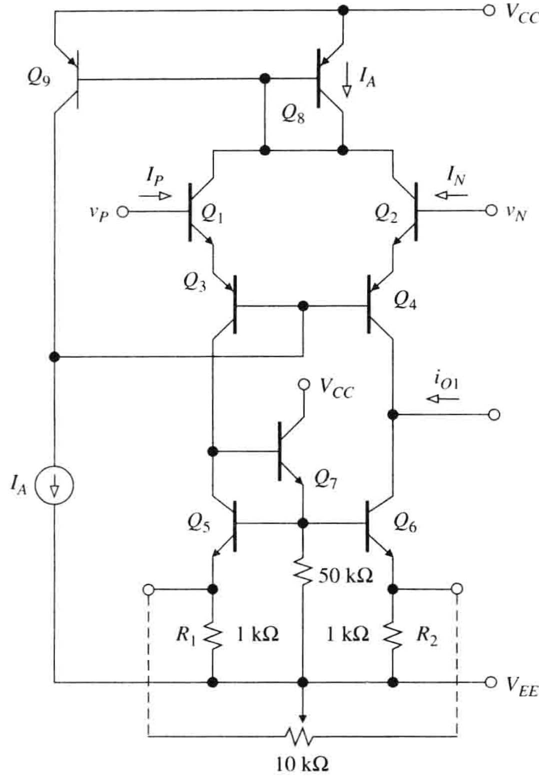


FIGURE 5.6
Detailed diagram of the input stage of the 741 op amp.

As we place the 741 in a circuit, its input transistors will draw I_P and I_N automatically from the surrounding elements. In fact, for the op amp to function, each input terminal must be provided with a series dc path through which current can flow (we have seen an example in connection with the GIC of Chapter 4). In the case of purely capacitive termination, the input current will charge or discharge the capacitor, making a periodic reinitialization necessary. Barring exceptions to be addressed in the next section, I_P and I_N flow *into* the op amp if its input transistors are *npn* BJTs or *p*-channel JFETs, and *out* of the op amp for *pnp* BJTs or *n*-channel JFETs.

Because of unavoidable mismatches between the two halves of the input stage, particularly between the β_{FS} of Q_1 and Q_2, I_P and I_N will themselves be mismatched. The average of the two currents is called the *input bias current,*

$$I_B = \frac{I_P + I_N}{2} \tag{5.8}$$

and their difference is called the *input offset current,*

$$I_{OS} = I_P - I_N \tag{5.9}$$

Usually I_{OS} is an order of magnitude smaller than I_B. While the polarity of I_B depends on the type of input transistors, that of I_{OS} depends on the direction of mismatch, so some samples of a given op amp family will have $I_{OS} > 0$, and others $I_{OS} < 0$.

Depending on the op amp type, I_B may range from nanoamperes to femtoamperes. The data sheets report typical as well as maximum values. For the 741C, which is the commercial version of the 741 family, the room-temperature ratings are: $I_B = 80$ nA typical, 500 nA maximum; $I_{OS} = 20$ nA typical, 200 nA maximum. For the 741E, which is the improved commercial version, $I_B = 30$ nA typical, 80 nA maximum; $I_{OS} = 3$ nA typical, 30 nA maximum. Both I_B and I_{OS} are temperature dependent, and these dependences are shown in Figs. 5A.8 and 5A.9, of Appendix 5A. The industry-standard OP77 op amp mentioned in Section 1.2, has $I_B = 1.2$ nA typical, 2.0 nA maximum; $I_{OS} = 0.3$ nA typical, 1.5 nA maximum.

Errors Caused by I_B and I_{OS}

A straightforward way of assessing the effect of the input currents is to find the output with all input signals set to zero. We shall illustrate two representative cases, namely, the cases of resistive and capacitive feedback shown in Fig. 5.7. Once we understand these cases, we can readily generalize to other circuits. Our analysis assumes that the op amp, aside from the presence of I_P and I_N, is ideal.

There are many circuits that, once their active inputs are set to zero, reduce to an equivalent circuit of the type of Fig. 5.7*a*, including the inverting and noninverting amplifiers, the summing and difference amplifiers, *I-V* converters, and others. By Ohm's law, the voltage at the noninverting input is $V_P = -R_p I_P$. Using the superposition principle, we have $v_O = (1+R_2/R_1)V_P + R_2 I_N = R_2 I_N - (1+R_2/R_1)R_p I_P$, or $v_O = E_O$, where

$$E_O = \left(1 + \frac{R_2}{R_1}\right)[(R_1 \parallel R_2)I_N - R_p I_P] \tag{5.10}$$

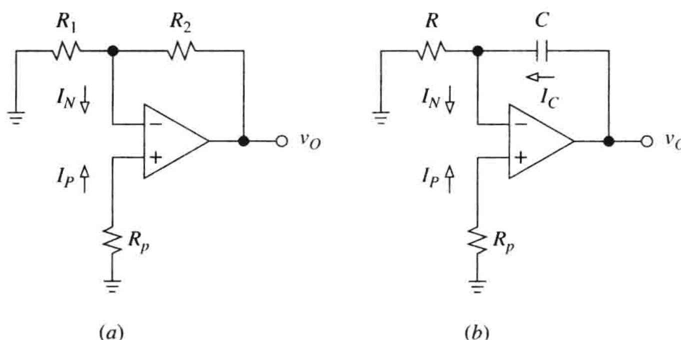


FIGURE 5.7
Estimating the output error due to the input bias currents for the case of (*a*) resistive, and (*b*) capacitive feedback.

This insightful form elicits a number of observations. First, in spite of the absence of any input signal, the circuit yields some output E_O. We regard this unwanted output as an error or, more properly, as *output dc noise*. Second, the circuit produces E_O by taking an input error, or *input dc noise,* and amplifying it by $(1 + R_2/R_1)$, which is aptly called the *dc noise gain* (recall that this gain is the reciprocal of the feedback factor β, hence the reason for calling $1/\beta$ the noise gain). Third, the input error consists of two terms, namely, the voltage drop $-R_p I_P$ due to I_P flowing out of R_p, and the term $(R_1 \| R_2)I_N$ arising as if I_N flew out of the parallel combination $R_1 \| R_2$ (R_1 and R_2 are *not* in parallel per se, but they only *appear* as such as they conspire with I_N to create the corresponding error term). Fourth, the two terms tend to compensate for each other since they have opposite polarities.

Depending on the application, the error E_O may be unacceptable and one must devise suitable means to reduce it to a tolerable level. Putting Eq. (5.10) in the form

$$E_O = \left(1 + \frac{R_2}{R_1}\right)\{[(R_1 \| R_2) - R_p]I_B - [(R_1 \| R_2) + R_p]I_{OS}/2\}$$

reveals that if we install a dummy resistance R_p, as shown, and we impose

$$R_p = R_1 \| R_2 \tag{5.11}$$

then the term involving I_B will be eliminated, leaving

$$E_O = \left(1 + \frac{R_2}{R_1}\right)(-R_1 \| R_2)I_{OS} \tag{5.12}$$

The error is now proportional to I_{OS}, which is typically an order of magnitude smaller than either I_P or I_N.

E_O can be reduced further by scaling down all resistances. For instance, reducing all resistance by a factor of 10 will leave gain unaffected, but will cause a tenfold reduction in the input error $-(R_1 \| R_2)I_{OS}$. Reducing resistances, however, increases power dissipation, so a compromise will have to be reached. If E_O is still unacceptable, selecting an op amp type with a lower I_{OS} rating is the next logical step. Other techniques for reducing E_O will be discussed in Section 5.6.

EXAMPLE 5.1. In the circuit of Fig. 5.7*a* let $R_1 = 22$ kΩ and $R_2 = 2.2$ MΩ, and let the op amp ratings be $I_B = 80$ nA and $I_{OS} = 20$ nA. (*a*) Calculate E_O for the case $R_p = 0$. (*b*) Repeat, but with $R_p = R_1 \| R_2$ in place. (*c*) Repeat part (*b*), but with all resistances simultaneously reduced by a factor of 10. (*d*) Repeat part (*c*), but with the op amp replaced by one with $I_{OS} = 3$ nA. Comment.

Solution.

(*a*) The dc noise gain is $1 + R_2/R_1 = 101$V/V; moreover, $(R_1 \| R_2) \cong 22$ kΩ. With $R_p = 0$, we have $E_O = 101 \times (R_1 \| R_2)I_N \cong 101 \times (R_1 \| R_2)I_B \cong 101 \times 22 \times 10^3 \times 80 \times 10^{-9} \cong 175$ mV.

(*b*) With $R_p = R_1 \| R_2 \cong 22$ kΩ in place, $E_O \cong 101 \times 22 \times 10^3 \times (\pm 20 \times 10^{-9}) = \pm 44$ mV, where we write "±" to reflect the fact that I_{OS} may be of either polarity.

(*c*) With $R_1 = 2.2$ kΩ, $R_2 = 220$ kΩ, and $R_p = 2.2$ kΩ, we get $E_O = 101 \times 2.2 \times 10^3 \times (\pm 20 \times 10^{-9}) \cong \pm 4.4$ mV.

(*d*) $E_O = 101 \times 2.2 \times 10^3 \times (\pm 3 \times 10^{-9}) \cong \pm 0.7$ mV. Summarizing, with R_p in place, E_O is reduced by 4; scaling the resistances reduces E_O by an additional factor of 10; finally, using a better op amp reduces it by yet another factor of 7.

Turning next to the circuit of Fig. 5.7*b*, we note that we still have $V_N = V_P = -R_p I_P$. Summing currents at the inverting-input node yields $V_N/R + I_N - I_C = 0$. Eliminating V_N, we get

$$I_C = \frac{1}{R}(RI_N - R_p I_P) = \frac{1}{R}[(R - R_p)I_B - (R + R_p)I_{OS}/2]$$

Applying the capacitance law $v = (1/C)\int i\,dt$, we readily get

$$v_O(t) = E_O(t) + v_O(0) \tag{5.13}$$

$$E_O(t) = \frac{1}{RC}\int_0^t [(R - R_p)I_B - (R + R_p)I_{OS}/2]\,d\xi \tag{5.14}$$

where $v_O(0)$ is the initial value of v_O. In the absence of any input signal, we expect the circuit to yield a constant output, or $v_O(t) = v_O(0)$. In practice, besides $v_O(0)$, it yields the *output error* $E_O(t)$, which is the result of integrating the *input error* $[(R - R_p)I_B - (R + R_p)I_{OS}/2]$ over time. Since I_B and I_{OS} are relatively constant, we can write $E_O(t) = [(R - R_p)I_B - (R + R_p)I_{OS}/2]t/RC$. The error is thus a voltage ramp, whose tendency is to drive the op amp into saturation.

It is apparent that installing a dummy resistance R_p such that

$$R_p = R \tag{5.15}$$

will reduce the error to

$$E_O(t) = \frac{1}{RC}\int_0^t -RI_{OS}\,d\xi \tag{5.16}$$

This error can be reduced further by component scaling, or by using an op amp with a lower I_{OS} rating.

EXAMPLE 5.2. In the circuit of Fig. 5.7*b* let $R = 100\ \text{k}\Omega$, $C = 1$ nF, and $v_O(0) = 0$ V. Assuming an op amp with $I_B = 80$ nA, $I_{OS} = 20$ nA, and $\pm V_{\text{sat}} = \pm 13$ V, find how long it takes for the op amp to enter saturation if (*a*) $R_p = 0$, and (*b*) $R_p = R$.

Solution.

(*a*) The input error is $RI_N \cong RI_B = 10^5 \times 80 \times 10^{-9} = 8$ mV. So, $v_O(t) = (RI_N/RC)t = 80t$, which represents a positive voltage ramp. Imposing $13 = 80t$ yields $t = 13/80 = 0.1625$ s.

(*b*) The input error is now $-RI_{OS} = \pm 2$ mV, indicating that the op amp may saturate at either rail. The time it takes to saturate is now extended in proportion to $0.1625 \times 80/20 = 0.65$ s.

Summarizing, to minimize the errors due to I_B and I_{OS}, adhere to the following rules whenever possible: (*a*) modify the circuit so that the resistances seen by I_P and I_N with all sources suppressed are equal, that is, impose $R_p = R_1 \parallel R_2$ in Fig. 5.7*a* and $R_p = R$ in Fig. 5.7*b*; (*b*) keep resistances as low as the application will allow; (*c*) use op amps with adequately low I_{OS} ratings.

5.3 LOW-INPUT-BIAS-CURRENT OP AMPS

Op amp designers strive to keep I_B and I_{OS} as small as other design constraints allow. Following are the most common techniques.

Superbeta-Input Op Amps

One way of achieving low I_Bs is by using input BJTs with extremely high current gains. Known as *superbeta transistors,* these BJTs achieve β_Fs in excess of 10^3 A/A by utilizing very thin base regions to minimize the recombination component[3] of the base current. This technique was pioneered with the LM308 op amp, whose input stage is shown in Fig. 5.8*a*. The heart of the circuit is the superbeta differential pair Q_1 and Q_2. These BJTs are connected in cascode with the standard-beta BJTs Q_3 and Q_4 to form a composite structure with high current gain as well as high breakdown voltage. Q_5 and Q_6 provide a bootstrapping function to bias Q_1 and Q_2 at zero base-collector voltage regardless of the input common-mode voltage. This avoids the low-breakdown limitations of the superbeta BJTs and also reduces collector-base leakage. Superbeta op amps have typically $I_B \cong 1$ nA or less.

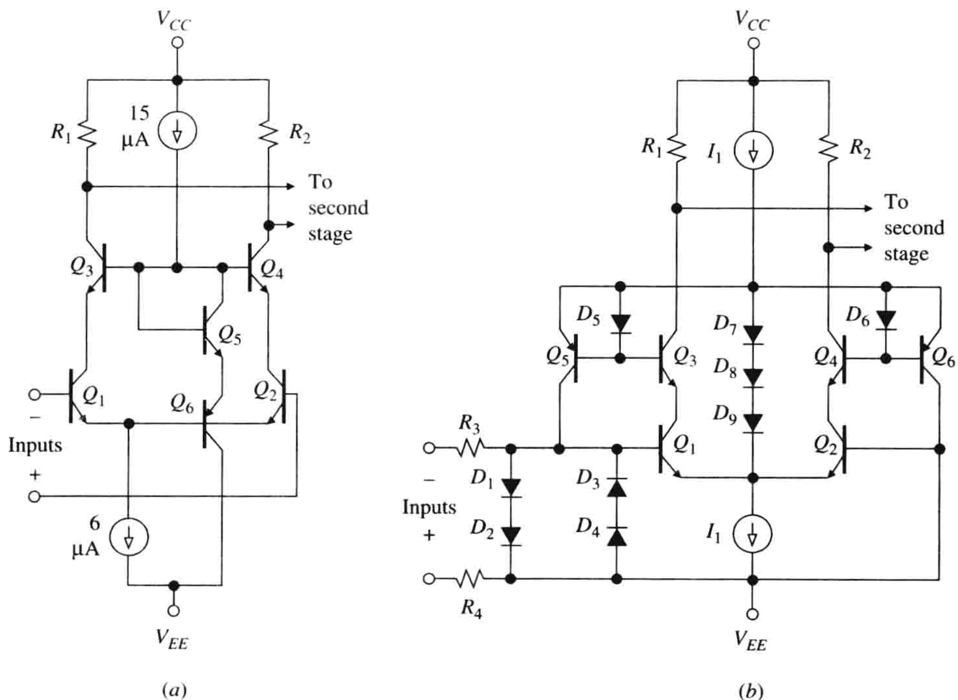


FIGURE 5.8
(*a*) Superbeta input stage, and (*b*) input-bias-current cancellation.

Input-Bias-Current Cancellation

Another popular technique for achieving low I_Bs is current cancellation.[3] Special circuitry anticipates the base currents needed to bias the input transistors, then itself supplies these currents internally, making the op amp appear to an outsider as if it were capable of operating without any input bias current.

Figure 5.8*b* shows the cancellation scheme utilized by the OP07 op amp. Once again, the heart of the circuit is the differential pair Q_1 and Q_2. The base currents of Q_1 and Q_2 are duplicated at the bases of common-base transistors Q_3 and Q_4, where they are sensed by current mirrors Q_5-D_5 and Q_6-D_6. The mirrors reflect these currents and then reinject them into the bases of Q_1 and Q_2, thus providing input-bias-current cancellation.

In practice, because of device mismatches, cancellation is not perfect, so the input pins will still draw residual currents. However, since these currents are now the result of a mismatch, they are typically an order of magnitude less than the actual base currents. We observe that I_P and I_N may flow either into or out of the op amp, depending on the direction of the mismatch. Moreover, I_{OS} is of the same order of magnitude as I_B, so there is no use installing a dummy resistance R_p in op amps with input-current cancellation. The OP07 ratings are $I_B = \pm 1$ nA and $I_{OS} = 0.4$ nA.

FET-Input Op Amps

As mentioned in Section 5.1, FET-input op amps generally exhibit much lower input bias currents than BJTs. We now wish to address this issue in more detail.

Consider first MOSFET op amps. The gate of a MOSFET forms a tiny capacitor with the body, so the gate of a well-fabricated MOSFET draws virtually no dc current. However, if the inputs are meant to be connected to external circuitry as in the case of general-purpose op amps, the fragile gates of the input FETs must be protected against electrostatic discharge (ESD) and electrical overstress (EOS). As depicted in Fig. 5.9*a*, the protective circuitry comprises internal diode clamps designed to prevent the gate voltages from raising a diode drop (0.7 V) above V_{DD} (D_H diodes)

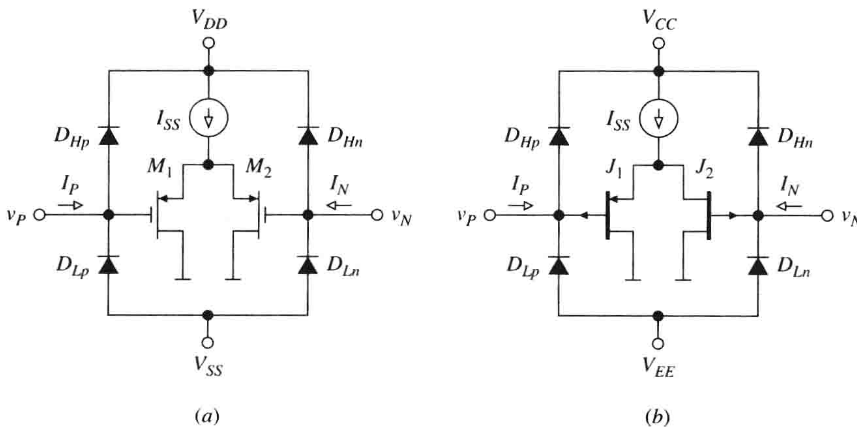


FIGURE 5.9
Input protection diodes for (*a*) CMOS op amps and (*b*) JFET-input op amp.

or dropping 0.7 V below V_{SS} (D_L diodes). In normal operation all diodes are reverse-biased, so each conducts a small reverse current I_R, which, at room temperature, is typically on the order of a few picoamperes. The data sheets usually give the input bias current with the common-mode input voltage $v_{CM} = (v_P + v_N)/2$ halfway between V_{DD} and V_{SS}. On average, this is a favorable situation because if the D_H and D_L diodes associated with the same input are matched, their I_Rs will cancel out, making I_P and I_N approach zero. However, raising v_{CM} above the cancellation level increases I_P and I_N because the amount of reverse bias increases for the D_L diodes and decreases for the D_H, making the I_R of the D_L bottom diodes prevail over that of the D_H diodes. Conversely, lowering v_{CM} below the cancellation level causes the I_R of the D_H diodes prevail over that of the D_L diodes, *reversing* the direction of I_P and I_N, whose magnitudes increase as v_{CM} is lowered further.

Figure 5.9*b* shows a similar diode network to protect the gates of the JFETs. However, in addition to protective-diode leakage, we now have also the leakage currents of the reverse-biased gate-channel junctions of the JFETs. These junctions are fabricated with significantly larger areas than the protective-diode junctions, so JFET leakage will prevail over diode leakage. The result is that I_P and I_N will flow *into* the op amp in the case of *p*-channel JFETs, as depicted in Fig. 5.9*b*, but *out* of the op amp if the differential input pair is implemented with *n*-channel JFETs.

To limit cluttering, the diode clamps were not shown explicitly in the circuits of Figs. 5.3 and 5.4. However, we need to be aware of their presence if we try to push v_P and/or v_N outside the supply voltages, for then the clamp diodes that will become conductive will cause I_P and/or I_N to shoot up exponentially.

Input-Bias-Current Drift

Figure 5.10 compares typical input-bias-current characteristics for different input-stage arrangements and technologies. We observe that in BJT-input devices, I_B tends to decrease with temperature, owing to the fact that β_F increases with temperature. However, for FET-input op amps, I_B increases exponentially with temperature. This is due to the fact that I_P and I_N are made up of I_Rs, and *I_R doubles for about every 10°C of temperature increase*. Extending this rule to I_B allows us to state that once we know I_B at some reference temperature T_0, we can estimate it at any other temperature T using

$$I_B(T) \cong I_B(T_0) \times 2^{(T-T_0)/10} \tag{5.17}$$

It is apparent from Fig. 5.10 that the low-input-bias-current advantages of FET-input op amps over their BJT-input counterparts tend to disappear at higher temperatures. Knowledge of the intended operating temperature range is an important factor when selecting the optimal device.

EXAMPLE 5.3. A certain FET-input op amp is rated at $I_B = 1$ pA at 25 °C. Estimate I_B at 100 °C.

Solution. $I_B(100\ °\mathrm{C}) \cong 10^{-12} \times 2^{(100-25)/10} = 0.18$ nA.

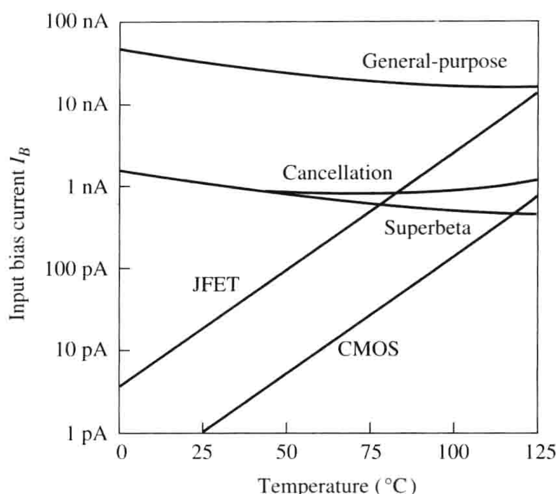


FIGURE 5.10
Typical input-bias-current characteristics.

Input Guarding

When applying op amps with ultralow input bias current, special attention must be paid to wiring and circuit construction in order to fully realize the capabilities of these devices. Data sheets usually provide helpful guidelines in this respect. Of special concern are leakage currents across the printed-circuit board. They can easily exceed I_B itself and thus defeat what has been so painstakingly achieved in terms of circuit design.

The effect of leakage can be reduced significantly by using guard rings around the input pins. As shown in Fig. 5.11, a guard consists of a conductive pattern held at the same potential as v_P and v_N. This pattern will absorb any leakages from other points on the board and thus prevent them from reaching the input pins. Guard rings also act as shields against noise pickup. For best results, board surfaces should be kept clean and moisture-free. If sockets are required, best results are obtained by using Teflon sockets or standoffs.

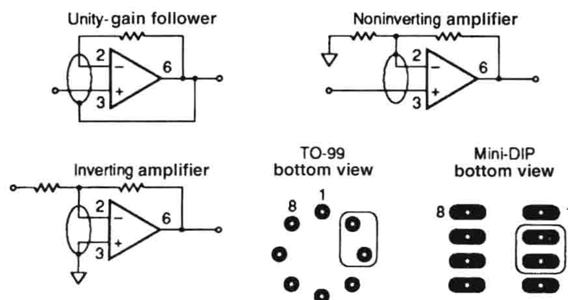


FIGURE 5.11
Guard-ring layout and connections.

5.4
INPUT OFFSET VOLTAGE

Shorting together the inputs of an op amp should yield $v_O = a(v_P - v_N) = a \times 0 = 0$ V. However, because of inherent mismatches between the input-stage halves processing v_P and v_N, a practical op amp will generally yield $v_O \neq 0$. To force v_O to zero, a suitable correction voltage must be applied between the input pins. This is tantamount to saying that the open-loop VTC does not go through the origin, but is shifted either to the left or to the right, depending on the direction of the mismatch. This shift is called the *input offset voltage* V_{OS}. As shown in Fig. 5.12, we can model a practical op amp with an ideal or offsetless op amp having a tiny source V_{OS} in series with one of its inputs. The VTC is now

$$v_O = a[v_P + V_{OS} - v_N] \tag{5.18}$$

To drive the output to zero, we need $v_P + V_{OS} - v_N = 0$, or

$$v_N = v_P + V_{OS} \tag{5.19}$$

Note that because of V_{OS}, we now have $v_N \neq v_P$.

As in the case of I_{OS}, the magnitude and polarity of V_{OS} varies from one sample to another of the same op amp family. Depending on the family, V_{OS} may range from millivolts to microvolts. The 741 data sheets give the following room-temperature ratings: for the 741C, $V_{OS} = 2$ mV typical, 6 mV maximum; and for the 741E, $V_{OS} = 0.8$ mV typical, 3 mV maximum. The OP77 ultralow offset voltage op amp has $V_{OS} = 10\ \mu$V typical, 50 μV maximum.

Errors Caused by V_{OS}

As in Section 5.2, we shall examine the effect of V_{OS} for the resistive-feedback and capacitive-feedback cases of Fig. 5.13. Note that we are omitting the dummy resistance R_P since the present analysis deliberately ignores I_B and I_{OS} to focus on V_{OS} alone. In Section 5.6 we shall address the general case in which I_B, I_{OS}, and V_{OS} are present simultaneously.

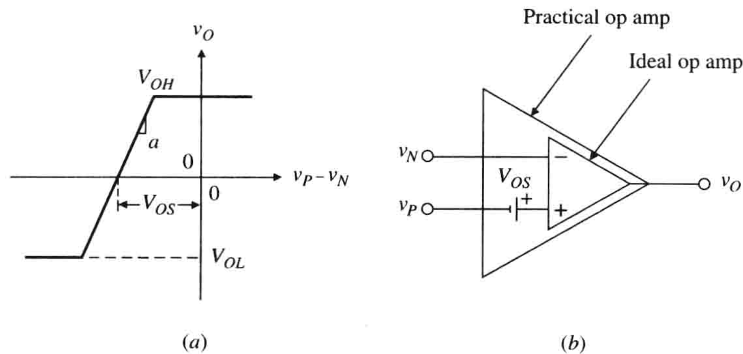


FIGURE 5.12
VTC and circuit model of an op amp with input offset voltage V_{OS}.

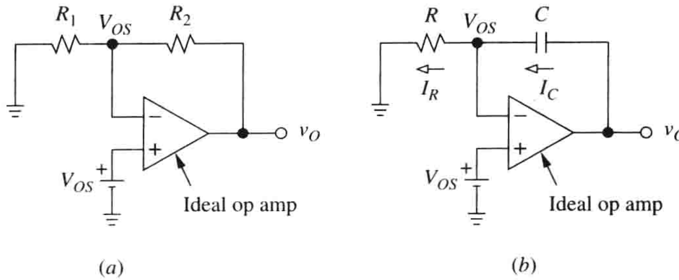


FIGURE 5.13
Estimating the output error due to V_{OS} for the case of (*a*) resistive, and (*b*) capacitive feedback.

In Fig. 5.13*a*, the offset-free op amp acts as a noninverting amplifier with respect to V_{OS}, so $v_O = E_O$, where

$$E_O = \left(1 + \frac{R_2}{R_1}\right) V_{OS} \tag{5.20}$$

is the output error, and $(1 + R_2/R_1)$ is again the dc noise gain. Clearly, the larger the noise gain, the larger the error. For instance, with $R_1 = R_2$, a 741C op amp yields $E_O = (1+1) \times (\pm 2 \text{ mV}) = \pm 4$ mV typical, $(1+1) \times (\pm 6 \text{ mV}) = \pm 12$ mV maximum. However, with $R_2 = 10^3 R_1$, it yields $E_O = (1 + 10^3) \times (\pm 2 \text{ mV}) \cong \pm 2$ V typical, ± 6 V maximum—quite an error! Conversely, we can use the present circuit to measure V_{OS}. For instance, let $R_1 = 10\ \Omega$ and $R_2 = 10\ \text{k}\Omega$, so that the dc noise gain is 1001 V/V and the combination $R_1 \parallel R_2$ is sufficiently small to make the effect of I_N negligible. Suppose we measure the output and find $E_O = -0.5$ V. Then, $V_{OS} \cong E_O/1001 \cong -0.5$ mV, a negative offset in this specific example.

In the circuit of Fig. 5.13*b* we note that since the offset-free op amp keeps $V_N = V_{OS}$, we have $I_C = I_R = V_{OS}/R$. Using again the capacitance law, we get $v_O(t) = E_O(t) + v_O(0)$, where the output error is now

$$E_O(t) = \frac{1}{RC} \int_0^t V_{OS}\, d\xi \tag{5.21}$$

or $E_O(t) = (V_{OS}/RC)t$. This voltage ramp, resulting from the integration of V_{OS} over time, tends, as we know, to drive the op amp into saturation.

Thermal Drift

Like most other parameters, V_{OS} is temperature-dependent, a feature expressed in terms of the *temperature coefficient*

$$\text{TC}(V_{OS}) = \frac{\partial V_{OS}}{\partial T} \tag{5.22}$$

where T is absolute temperature, in kelvins, and $\text{TC}(V_{OS})$ is in microvolts per degree Celsius. For low-cost, general-purpose op amps such as the 741, $\text{TC}(V_{OS})$ is typically

on the order of 5 μV/°C. Thermal drift stems from inherent mismatches as well as thermal gradients across the two halves of the input stage. Op amps specifically designed for low-input offset also tend to exhibit lower thermal drifts, thanks to superior matching and thermal tracking at the input stage. The OP77 has $\text{TC}(V_{OS}) =$ 0.1 μV/°C typical, 0.3 μV/°C maximum.

Using the average value of the temperature coefficient, one can estimate V_{OS} at a temperature other than 25 °C as

$$V_{OS}(T) \cong V_{OS}(25\ ^\circ\text{C}) + \text{TC}(V_{OS})_{\text{avg}} \times (T - 25\ ^\circ\text{C}) \tag{5.23}$$

For instance, an op amp with $V_{OS}(25\ ^\circ\text{C}) = 1$ mV and $\text{TC}(V_{OS})_{\text{avg}} = 5\ \mu\text{V}/^\circ\text{C}$ would have $V_{OS}(70\ ^\circ\text{C}) = 1\ \text{mV} + (5\ \mu\text{V}) \times (70 - 25) = 1.225$ mV.

Common-Mode Rejection Ratio (CMRR)

In the absence of input offset, an op amp should respond only to the voltage difference between its inputs, or $v_O = a(v_P - v_N)$. A practical op amp is somewhat sensitive also to the common-mode input voltage $v_{\text{CM}} = (v_P + v_N)/2$. Its transfer characteristic is thus $v_O = a(v_P - v_N) + a_{\text{cm}} v_{\text{CM}}$, where a is the differential-mode gain, and a_{cm} is the common-mode gain. Rewriting as $v_O = a[v_P - v_N + (a_{\text{cm}}/a)v_{\text{CM}}]$, and recalling that the ratio a/a_{cm} is the common-mode rejection ratio CMRR, we have

$$v_O = a\left(v_P + \frac{v_{\text{CM}}}{\text{CMRR}} - v_N\right)$$

Comparison with Eq. (5.18) indicates that the sensitivity to v_{CM} can be modeled with an input-offset-voltage term of value $v_{\text{CM}}/\text{CMRR}$. The common-mode sensitivity stems from the fact that a change in v_{CM} will alter the operating points of the input-stage transistors and cause a change at the output. It is comforting to know that such a complex phenomenon can be reflected to the input in the form of a mere offset error! We thus redefine the CMRR as

$$\frac{1}{\text{CMRR}} = \frac{\partial V_{OS}}{\partial v_{\text{CM}}} \tag{5.24}$$

and interpret it as the change in V_{OS} brought about by a 1-V change in v_{CM}. We express 1/CMRR in microvolts per volt. Because of stray capacitances, the CMRR deteriorates with frequency. Typically, it is high from dc to a few tens or a few hundreds of hertz, after which it rolls off with frequency at the rate of −20 dB/dec.

Data sheets usually give CMRR in decibels. As we know, the conversion to microvolts per volt is readily accomplished via

$$\frac{1}{\text{CMRR}} = 10^{-\text{CMRR}_{\text{dB}}/20} \tag{5.25}$$

where CMRR_{dB} represents the decibel value of CMRR. From Fig. 5A.4, the dc ratings for the 741 op amp are $\text{CMRR}_{\text{dB}} = 90$ dB typical, 70 dB minimum, indicating that V_{OS} changes with v_{CM} at the rate of $1/\text{CMRR} = 10^{-90/20} = 31.6\ \mu\text{V/V}$ typical,

and $10^{-70/20} = 316\,\mu$V/V maximum. The OP77 op amp has $1/\text{CMRR} = 0.1\mu$ V/V typical, 1 μV/V maximum. Figure 5A.6 shows that the CMRR of the 741 starts to roll off just above 100 Hz.

Since op amps keep v_N fairly close to v_P, we can write $v_{CM} \cong v_P$. The CMRR is of no concern in inverting applications, where $v_P = 0$. However, it may pose problems when v_P is allowed to swing, as in an instrumentation amplifier.

EXAMPLE 5.4. The difference amplifier of Fig. 2.13*a* uses a 741 op amp and a perfectly matched resistance set with $R_1 = 10$ kΩ and $R_2 = 100$ kΩ. Suppose the inputs are tied together and driven with a common signal v_I. Estimate the typical change in v_O if (*a*) v_I is slowly changed from 0 to 10 V, and (*b*) v_I is a 10-kHz, 10-V peak-to-peak sine wave.

Solution.

(*a*) At dc we have $1/\text{CMRR} = 10^{-90/20} = 31.6\ \mu$V/V, typical. The common-mode change at the op amp input pins is $\Delta v_P = [R_2/(R_1 + R_2)]\Delta v_I = [100/(10 + 100)]10 = 9.09$ V. Thus, $\Delta V_{OS} = (1/\text{CMRR})\Delta v_P = 31.6 \times 9.09 = 287\ \mu$V. The dc noise gain is $1 + R_2/R_1 = 11$ V/V. Hence, $\Delta v_O = 11 \times 287 = 3.16$ mV.

(*b*) From the CMRR curve of Fig. 5A.6 we find $\text{CMRR}_{\text{dB}}(10\text{ kHz}) \cong 57$ dB. So, $1/\text{CMRR} = 10^{-57/20} = 1.41$ mV/V, $\Delta V_{OS} = 1.41 \times 9.09 = 12.8$ mV (peak-to-peak), and $\Delta v_O = 11 \times 12.8 = 0.141$ V (peak-to-peak). The output error at 10 kHz is much worse than at dc.

Power-Supply Rejection Ratio (PSRR)

If we change one of the op amp supply voltages V_S by a given amount ΔV_S, the operating points of the internal transistors will be altered, generally causing a small change in v_O. By analogy with the CMRR, we model this phenomenon with a change in the input offset voltage, which we express in terms of the *power-supply rejection ratio* (PSRR) as $(1/\text{PSRR}) \times \Delta V_S$. The parameter

$$\frac{1}{\text{PSRR}} = \frac{\partial V_{OS}}{\partial V_S} \tag{5.26}$$

represents the change in V_{OS} brought about by a 1-V change in V_S, and is expressed in microvolts per volt. Like the CMRR, the PSRR deteriorates with frequency.

Some data sheets give separate PSRR ratings, one for changes in V_{CC} and the other for changes in V_{EE}. Others specify the PSRR for V_{CC} and V_{EE} changing symmetrically. The PSRR_{dB} ratings of most op amps fall in the range of 80 dB to 120 dB. The devices of superior matching usually offer the highest PSRRs. From Fig. 5A.4, the 1/PSRR ratings for the 741C, which are given for symmetric supply changes, are 30 μV/V typical, 150 μV/V maximum. This means that changing, for instance, the supply voltages from ± 15 V to ± 12 V yields $\Delta V_{OS} = (1/\text{PSRR})\Delta V_S = (30\ \mu\text{V})(15 - 12) = \pm 90\ \mu$V typical, $\pm 450\ \mu$V maximum. The OP77 op amp has $1/\text{PSRR} = 0.7\ \mu$V/V typical, 3 μV/V maximum.

When the op amp is powered from well-regulated and properly bypassed supplies, the effect of the PSRR is usually negligible. Otherwise, any variation on the supply busses will induce a corresponding variation in V_{OS}, which in turn is amplified by the noise gain. A classic example is offered by audio preamplifiers, where the residual 60 Hz (or 120 Hz) ripple on the supply rails may cause intolerable hum

at the output. Another case in point is offered by switchmode power supplies, whose high-frequency ripple is usually inadequately rejected by op amps, indicating that these supplies are unsuited to high-precision analog circuitry.

EXAMPLE 5.5. A 741 op amp is connected as in Fig. 5.13*a* with $R_1 = 100\ \Omega$ and $R_2 = 100\ \text{k}\Omega$. Predict the typical as well as the maximum ripple at the output for a power-supply ripple of 0.1 V (peak-to-peak) at 120 Hz.

Solution. The 741 data sheets do not show the PSRR rolloff with frequency, so let us use the ratings given at dc, keeping in mind that the results will be optimistic. The induced ripple at the input is $\Delta V_{OS} = (30\ \mu\text{V})0.1 = 3\ \mu\text{V}$ typical, 15 μV maximum (peak-to-peak). The noise gain is $1 + R_2/R_1 \cong 1000$ V/V, so the output ripple is $\Delta v_O =$ 3 mV typical, 15 mV maximum (peak-to-peak).

Change of V_{OS} with the Output Swing

In a practical op amp the open-loop gain a is finite, so the difference $v_P - v_N$ changes also with the output swing Δv_O by the amount $\Delta v_O/a$. This effect can conveniently be regarded as an effective offset voltage change $\Delta V_{OS} = \Delta v_O/a$. Even an op amp with $V_{OS} = 0$ for $v_O = 0$ will exhibit some input offset for $v_O \neq 0$. To illustrate with a numerical example, consider the inverting amplifier of Fig. 5.14*a*, designed to provide a closed-loop gain of $-18/2 = -9$ V/V using an op amp with $a = 10^4$ V/V. This circuit has $\beta = 0.1$, $T = 10^3$, and $A = -9/(1 + 1/10^3) = -8.991$ V/V. With $v_I = 1.0$ V the op amp gives $v_O = Av_I = -8.991$ V, and to sustain this voltage it needs $v_N = -v_O/a = +0.8991$ mV. Alternatively, we can regard the op amp as ideal ($a = \infty$) but afflicted by an input offset voltage $V_{OS} = 0.8991$ mV to account for the closed-loop gain error due to finite a. This alternative viewpoint is depicted in Fig. 5.14*b*. The two viewpoints are contrasted mathematically as

$$v_O = A_{\text{ideal}} \frac{1}{1 + 1/(a\beta)} v_I \qquad v_O = A_{\text{ideal}} v_I + \frac{1}{\beta} V_{OS}$$

Note how the second viewpoint differentiates between the signal gain A_{ideal} and the noise gain $1/\beta$. It also provides yet another insightful interpretation of the gain error, namely, a form of input noise V_{OS}.

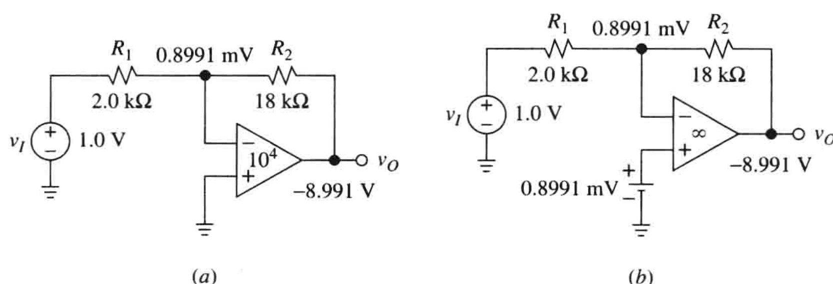


FIGURE 5.14
Using the input offset voltage to model the effect of finite open-loop gain a.

Complete Expression for the Input Offset Voltage V_{OS}

We summarize this section with a comprehensive expression for V_{OS},

$$V_{OS} = V_{OS0} + \mathrm{TC}(V_{OS})\Delta T + \frac{\Delta v_P}{\mathrm{CMRR}} + \frac{\Delta V_S}{\mathrm{PSRR}} + \frac{\Delta v_O}{a} \tag{5.27}$$

where V_{OS0}, the *initial input offset voltage,* is the value of V_{OS} at some reference operating point, such as ambient temperature, nominal supply voltages, and v_P and v_O halfway between the supply voltages. This parameter itself drifts with time. As an example, the OP77 has a long-term stability of 0.2 μV/month. In error-budget analysis, the various offset changes are combined *additively* when we wish to estimate the *worst-case* change, and in *root-sum-square (rss)* fashion when we are interested in the *most probable* change. Looking back at Eq. (5.27), we find it remarkable that all the op amp imperfections considered thus far lend themselves to be modeled via mere V_{OS} components regardless of their complexity. Also, by reducing them to a common form, we can easily compare them against each other, and decide what measures to take if we wish to reduce the intolerably large ones.

EXAMPLE 5.6. An op amp has the following ratings: $a = 10^5$ V/V typical, 10^4 V/V minimum, $\mathrm{TC}(V_{OS})_{\mathrm{avg}} = 3\ \mu$V/°C, and $\mathrm{CMRR_{dB}} = \mathrm{PSRR_{dB}} = 100$ dB typical, 80 dB minimum. Estimate the worst-case as well as the most probable change in V_{OS} over the following range of operating conditions: $0\ °\mathrm{C} \le T \le 70\ °\mathrm{C}$, $V_S = \pm 7.5\ \mathrm{V} \pm 10\%$, $-1\ \mathrm{V} \le v_P \le +1$ V, and $-5\ \mathrm{V} \le v_O \le +5$ V.

Solution. The thermal change from room temperature is $\Delta V_{OS1} = (3\ \mu\mathrm{V/°C})(70 - 25)\mathrm{°C} = 135\ \mu$V. With $1/\mathrm{CMRR} = 1/\mathrm{PSRR} = 10^{-100/20} = 10\ \mu$V/V typical, 100 μV/V maximum, the changes with v_P and V_S are $\Delta V_{OS2} = (\pm 1\ \mathrm{V})/\mathrm{CMRR} = \pm 10\ \mu$V typical, $\pm 100\ \mu$V maximum; $\Delta V_{OS3} = 2 \times (\pm 0.75\ \mathrm{V})/\mathrm{PSRR} = \pm 15\ \mu$V typical, $\pm 150\ \mu$V maximum. Finally, the change with v_O is $\Delta V_{OS4} = (\pm 5\ \mathrm{V})/a = \pm 50\ \mu$V typical, $\pm 500\ \mu$V maximum. The worst-case change in V_{OS} is $\pm(135 + 100 + 150 + 500) = \pm 885\ \mu$V. The most probable change is $\pm(135^2 + 10^2 + 15^2 + 50^2)^{1/2} = \pm 145\ \mu$V.

5.5 LOW-INPUT-OFFSET-VOLTAGE OP AMPS

The initial input offset voltage of an op amp is due primarily to mismatches in the transistors of its input stage. Here we shall consider two representative cases, the bipolar op amp of Fig. 5.1 and the CMOS op amp of Fig. 5.4*a*, but without dwelling on the detailed derivations, which are beyond our scope here.

Bipolar Input Offset Voltage

Mismatches are random phenomena, so we are interested in the most probable input offset voltage of the Q_1-Q_2-Q_3-Q_4 structure of Fig. 5.1. This takes on the form[3]

$$V_{OS(\mathrm{BJT})} \cong V_T \sqrt{\left(\frac{\Delta I_{sp}}{I_{sp}}\right)^2 + \left(\frac{\Delta I_{sn}}{I_{sn}}\right)^2} \tag{5.28}$$

where $V_T = kT/q$ is the *thermal voltage*, proportional to absolute temperature $T(V_T \cong 26$ mV at room temperature); I_{sn} and I_{sp} are the *collector saturation currents* appearing in the BJT i-v characteristics, $I_{C(npn)} = I_{sn}\exp(V_{BE}/V_T)$ and $I_{C(pnp)} = I_{sp}\exp(V_{EB}/V_T)$; the ratios $\Delta I_{sp}/I_{sp}$ and $\Delta I_{sn}/I_{sn}$ represent fractional variation of I_{sp} and I_s. BJT saturation currents take on the common form[4]

$$I_s = \frac{qD_B}{N_B} \times n_i^2(T) \times \frac{A_E}{W_B} \tag{5.29}$$

where D_B and N_B are the minority-carrier diffusion constant and doping concentration in the base region, $n_i^2(T)$ is the intrinsic carrier concentration, a strong function of T, and A_E and W_B are the emitter-junction area and the base-region width. It is apparent that in the case of perfectly matched BJT pairs we have $\Delta I_{sn} = 0$ and $\Delta I_{sp} = 0$, so $V_{OS} = 0$. On the other hand, with typical mismatches of, say, 5% we have $V_{OS} = (26\text{ mV})(0.05^2 + 0.05^2)^{1/2} = 1.84$ mV at room temperature. The thermal drift is $\text{TC}(V_{OS}) = \partial V_{OS}/\partial T = k/q$, or

$$\text{TC}(V_{OS}) = \frac{V_{OS}}{T} \tag{5.30}$$

indicating that at room temperature ($T = 300$ K) the bipolar structure exhibits a $\text{TC}(V_{OS})$ of 3.3 μV/°C for every millivolt of offset voltage.

CMOS Offset Voltage

The most probable input offset voltage of the M_1-M_2-M_3-M_4 stage of Fig. 5.4*a* takes on the form[4]

$$V_{OS(\text{CMOS})} \cong \frac{V_{OVp}}{2}\sqrt{\left(\frac{\Delta k_n}{k_n}\right)^2 + \left(\frac{\Delta k_p}{k_p}\right)^2 + \left(\frac{\Delta V_{tn}}{0.5V_{OVp}}\right)^2 + \left(\frac{\Delta V_{tp}}{0.5V_{OVp}}\right)^2} \tag{5.31}$$

where V_{OVp} is the overdrive voltage of the pMOSFETs; V_{tn} and V_{tp} are the threshold voltages of the nMOSFETs and pMOSFETs; k_n and k_p are the device transconductance parameters appearing in the MOSFET i-v characteristics, $I_{D(n\text{MOSFET})} = (k_n/2)V_{OVn}^2$ and $I_{D(p\text{MOSFET})} = (k_p/2)V_{OVp}^2$. These parameters are

$$k_n = \mu_n\frac{\varepsilon_{ox}}{t_{ox}}\frac{W_n}{L_n} \qquad k_p = \mu_p\frac{\varepsilon_{ox}}{t_{ox}}\frac{W_p}{L_p} \tag{5.32}$$

Here, μ_n and μ_p are the electron and hole mobilities, ε_{ox} and t_{ox} are the permittivity and the thickness of the gate-body oxide layer, and W_n/L_n and W_p/L_p are the channel width/length ratios of the nMOSFET and the pMOSFET.

Matching Considerations

Mismatches in integrated circuits stem from fluctuations in *doping* and *device dimensions*, as well as mechanical stress and others. Doping fluctuations affect I_{sn} and

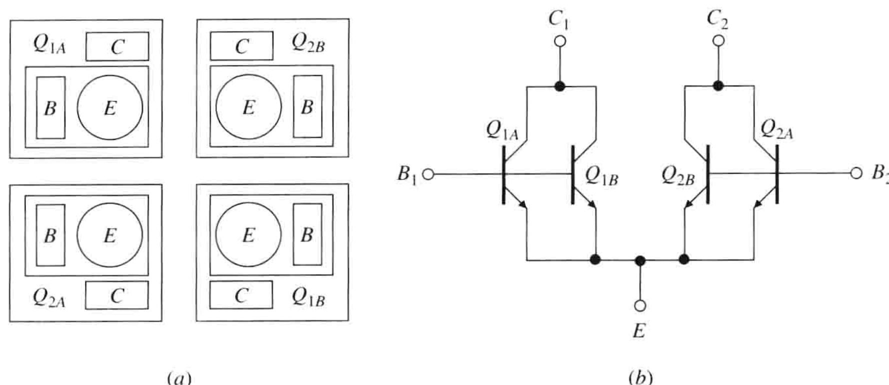


FIGURE 5.15
Common-centroid topology: (*a*) layout and (*b*) interconnections.

I_{sp} (via N_B), and V_{tn} and V_{tp}. Dimension fluctuations affect A_E/W_B ratios, W/L ratios, and the oxide thickness t_{ox}. A common way of reducing the sensitivity to doping irregularities and edge resolution is by fabricating input-stage BJTs with large emitter areas A_E, and input-stage MOSFETs with large channel dimensions W and L (in CMOS op amps, large transistor sizes result also in better noise performance, a topic addressed in Chapter 7).

Another important form of mismatch stems from *thermal gradients* across the chip, which affect V_{BE} and V_{EB} in BJTs, and V_{tn} and V_{tp} in MOSFETs. It is worth remembering the following thermal coefficients:

$$\text{TC}(V_{BE}) \cong \text{TC}(V_{EB}) \cong -2\text{ mV/}^\circ\text{C} \qquad \text{TC}(V_{tp}) \cong -\text{TC}(V_{tn}) \cong +2\text{ mV/}^\circ\text{C}$$

This means that a temperature difference of just 1°C between the transistors of a differential pair would contribute 2 mV to V_{OS}! The input-stage sensitivity to thermal gradients is reduced by a symmetrical device placement technique known as *common-centroid layout*.[2,3] As exemplified in Fig. 5.15 for the case of a bipolar input pair, each transistor is made up of two identical halves connected in parallel, but laid out diagonally opposite to each other. The resulting quad structure provides a multifold symmetry that tends to cancel out the effects of gradient-induced mismatches.

Offset Voltage Trimming

Op amp manufacturers reduce V_{OS} further via on-chip trimming techniques.[2]

A popular class involves the physical modification of one of the two halves making up the input stage, either by means of a laser trim, or by selectively shorting or opening suitable trimming links. The concept is illustrated in Fig. 5.16 for the case of the industry-standard OP07 op amp, which pioneered this technique (for more details, search "OP07" on the web). Each collector resistor is made up of a fixed part R_c in series with a string of smaller, binary-weighted resistors, each having a reverse-biased junction in parallel with it. During the wafer-probing stage the inputs are shorted together, the output is measured, and a suitable algorithm specifies which

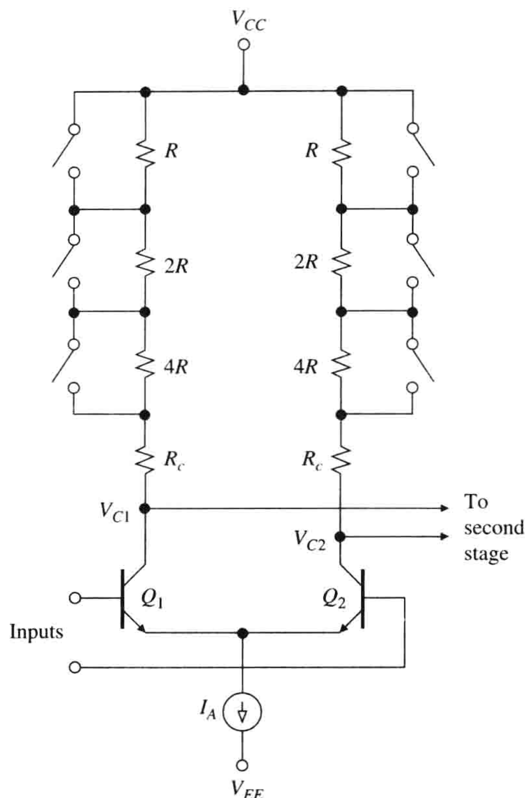


FIGURE 5.16
On-chip V_{OS} trimming using shortable links.

links must be closed to drive the output close to zero. For instance, if a mismatch between Q_1 and Q_2 results in $V_{C1} < V_{C2}$, we need to selectively short some of Q_1's resistors in order to raise V_{C1} until it approaches V_{C2}, giving the appearance of offsetless behavior. A switch is closed by passing a large reverse current through the corresponding diode, which then becomes a short circuit. Thanks to this technique, also known as *zener zapping*, the OP07E version offers $V_{OS} = 30\ \mu$V and TC(V_{OS}) $= 0.3\ \mu$V/°C.

A variant of this scheme uses aluminum fuse links to form initially closed switches. During the trim phase the switches are selectively opened via suitable current pulses.

Another class of trimming schemes, popular especially with CMOS op amps, uses on-chip nonvolatile memory to store trim data that are then converted to suitable adjustment currents via on-chip D/A converters.[2]

Autozero and Chopper-Stabilized Op Amps

On-chip trimming nulls V_{OS} at a specific set of environmental and operating conditions. As these conditions change, so does V_{OS}. To meet the stringent requirements

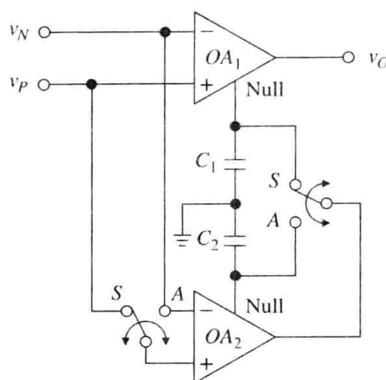


FIGURE 5.17
Chopper-stabilized op amp (CSOA).

of high-precision applications, special techniques have been developed to effectively reduce the input offset as well as low-frequency noise even further. Two popular such methods are the *autozero* (AZ) and *chopper stabilization* (CS) techniques. The AZ technique is a *sampling* technique[6] that samples the offset and low-frequency noise and then subtracts it from the contaminated signal to give offset-free appearance. The CS technique is a *modulation* technique[6] that modulates the input signal to a higher frequency where there is no dc offset or low-frequency noise, and then demodulates the amplified signal thus stripped of offset and low-frequency errors back to the baseband.

Figure 5.17 illustrates the AZ principle for the case of the ICL7650S op amp, the first popular op amp to realize this technique in monolithic form. The heart of the device is OA_1, a conventional, high-speed amplifier referred to as the *main amplifier.* A second amplifier, called the *nulling amplifier* and denoted as OA_2, continuously monitors OA_1's input offset error V_{OS1} and drives it to zero by applying a suitable correcting voltage at OA_1's null pin. This mode of operation is called the *sampling mode.*

Note, however, that OA_2 too has an input offset V_{OS2}, so it must correct its own error before attempting to improve OA_2's error. This is achieved by momentarily disconnecting OA_2 from the main amplifier, shorting its inputs together, and coupling its output to its own null pin. This mode, referred to as the *autozero mode,* is activated by flipping the MOS switches from the S (sampling) position to the A (autozero) position. During the autozero mode, the correction voltage for OA_1 is momentarily held by C_1, which therefore acts as an analog memory for this voltage. Similarly, C_2 holds the correction voltage for OA_2 during the sampling mode.

Alternation between the two modes takes place at a typical rate of a few hundred cycles per second, and is controlled by an on-chip oscillator, making the AZ operation transparent to the user. The error-holding capacitors (0.1 μF for the aforementioned ICL7650S) are supplied off-chip by the user. The room-temperature rating for the ICL7650S is $V_{OS} = \pm 0.7\ \mu$V.

Like AZ op amps, CS op amps also utilize a pair of capacitors to realize the modulation/demodulation function. In some devices these capacitors are encapsulated

in the IC package itself to save space. Examples of this type of CS op amp are the LTC1050 with $V_{OS} = 0.5\ \mu V$ and $TC(V_{OS}) = 0.01\ \mu V/^\circ C$ typical, and the MAX420 with $V_{OS} = 1\ \mu V$ and $TC(V_{OS}) = 0.02\ \mu V/^\circ C$.

The impressive dc specifications of AZ and CS op amps do not come for free, however. Since the nulling circuit is a sampled-data system, clock-feedthrough noise and frequency aliasing problems arise, which need be taken into consideration when selecting the device best suited to the application.

AZ and CS op amps can be used either alone or as part of composite amplifiers to improve existing input specifications.[7] To fully realize these specifications, considerable attention must be paid to circuit board layout and construction.[7] Of particular concern are input leakage currents and thermocouple effects arising at the junction of dissimilar metals. They can grossly degrade the input specifications of the device and completely defeat what has been so painstakingly achieved in terms of circuit design. Consult the data sheets for valuable hints in this regard.

5.6 INPUT OFFSET ERROR AND COMPENSATION TECHNIQUES

We are now ready to investigate the effect of I_{OS} and V_{OS} acting simultaneously. We begin with the familiar amplifiers of Fig. 5.18 (ignore the 10-kΩ potentiometers for the time being).

Using Eqs. (5.12) and (5.20), along with the superposition principle, it is readily seen that both circuits yield

$$v_O = A_s v_I + E_O \tag{5.33a}$$

$$E_O = \left(1 + \frac{R_2}{R_1}\right)[V_{OS} - (R_1 \parallel R_2)I_{OS}] = \frac{1}{\beta}E_I \tag{5.33b}$$

where $A_s = -R_2/R_1$ for the inverting amplifier, and $A_s = 1 + R_2/R_1$ for the noninverting one. We call A_s the *signal gain* to distinguish it from the *dc noise gain,* which is $1/\beta = 1 + R_2/R_1$ for *both* circuits. Moreover, $E_I = V_{OS} - (R_1 \parallel R_2)I_{OS}$ is the *total offset error referred to the input,* and E_O the *total offset error referred to the output.* The negative sign does not necessarily imply a tendency by

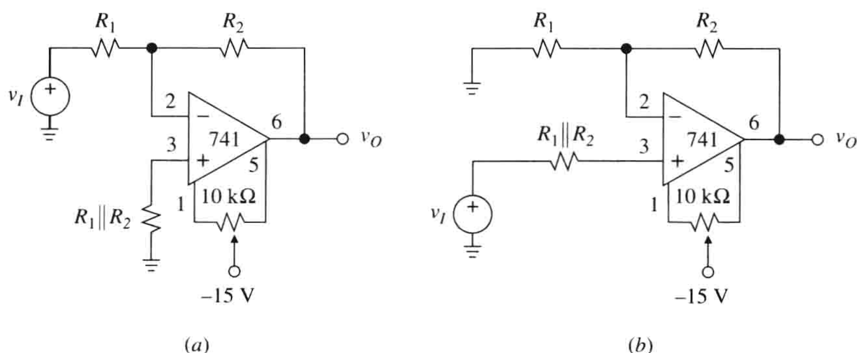


FIGURE 5.18
(*a*) Inverting, and (*b*) noninverting amplifiers with internal offset-error nulling.

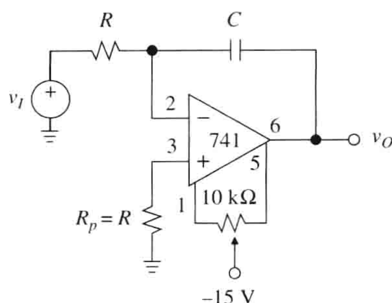


FIGURE 5.19
Integrator with internal offset-error nulling.

the two terms to compensate for each other, since V_{OS} and I_{OS} may be of either polarity. A prudent designer will take a conservative viewpoint and combine them additively.

The presence of the output error E_O may or may not be a drawback, depending on the application. In audio applications, where dc voltages are usually blocked out through capacitive coupling, offset voltages are seldom of major concern. Not so in low-level signal detection, such as thermocouple or strain-gauge amplification, or in wide dynamic-range applications, such as logarithmic compression and high-resolution data conversion. Here v_I may be of comparable magnitude to E_I, so its information content may easily be obliterated. The problem then arises of reducing E_I below a tolerable level.

Turning next to the integrator of Fig. 5.19, we use Eqs. (5.16) and (5.21) and the superposition principle to write

$$v_O(t) = -\frac{1}{RC}\int_0^t [v_I(\xi) + E_I]\, d\xi + v_O(0) \tag{5.34a}$$

$$E_I = RI_{OS} - V_{OS} \tag{5.34b}$$

Now the effect of V_{OS} and I_{OS} is to offset v_I by the error E_I. Even with $v_I = 0$, the output will ramp up or down until saturation is reached.

The input-referred error E_I in Eqs. (5.33*b*) and (5.34*b*) can be nulled by means of a suitable trimmer, as we are about to see. However, as we know, trimmers increase production costs and drift with temperature and time. A wise designer will try minimizing E_I by a combination of circuit tricks, such as resistance scaling and op amp selection. Only as a last resort should one turn to trimmers. Offset nulling techniques are classified as *internal* and *external.*

Internal Offset Nulling

Internal nulling is based on the deliberate unbalancing of the input stage to make up for inherent mismatches and drive the error to zero. This imbalance is introduced by

means of an external trimmer, as recommended in the data sheets. Figure 5.6 shows the trimmer connection for the internal nulling for the 741 op amp. The input stage consists of two nominally identical halves: the Q_1-Q_3-Q_5-R_1 half to process v_P and the Q_2-Q_4-Q_6-R_2 half to process v_N. Varying the wiper away from its center position will place more resistance in parallel with one side and less with the other, thus unbalancing the circuit. To calibrate the amplifiers of Fig. 5.18, we set $v_I = 0$ and we adjust the wiper for $v_O = 0$. To calibrate the integrator of Fig. 5.19, we set $v_I = 0$ and we adjust the wiper for v_O as steady as possible in the vicinity of 0 V.

From the 741C data sheets of Fig. 5A.3, we note that the *offset-voltage adjustment range* is typically ± 15 mV, indicating that for this compensating scheme to succeed we must have $|E_I| < 15$ mV. Since the 741C has $V_{OS} = 6$ mV maximum, this leaves 9 mV for the offset term due to I_{OS}. If this term exceeds 9 mV, we must either scale down the external resistances or resort to external nulling, to be discussed below.

EXAMPLE 5.7. A 741C op amp is to be used in the circuit of Fig. 5.18*a* to yield $A_s = -10$ V/V. Specify suitable resistances that will maximize the input resistance R_i of the circuit.

Solution. Since $R_i = R_1$, we need to maximize R_1. Imposing $R_2 = 10R_1$, and $V_{OS(\max)} + (R_1 \parallel R_2)I_{OS(\max)} \leq 15$ mV, we get $R_1 \parallel R_2 \leq (15 \text{ mV} - 6 \text{ mV})/(200 \text{ nA}) = 45$ kΩ, or $1/R_1 + 1/(10R_1) \geq 1/(45 \text{ k}\Omega)$. Solving yields $R_1 \leq 49.5$ kΩ. Use the standard values $R_1 = 47$ kΩ, $R_2 = 470$ kΩ, and $R_p = 43$ kΩ.

Internal offset nulling can be applied to any of the circuits studied so far. In general, the nulling scheme varies from one op amp family to another. To find the recommended nulling scheme for a given device, consult the data sheets. We observe that dual- and quad-op-amp packages usually do not have provisions for internal nulling due to lack of available pins.

External Offset Nulling

External nulling is based on the injection of an adjustable voltage or current into the circuit to compensate for its offset error. This scheme does not introduce any additional imbalances in the input stage, so there is no degradation in drift, CMRR, or PSRR.

The most convenient point of injection of the correcting signal depends on the particular circuit. For inverting-type configurations like the amplifier and integrator of Fig. 5.20, we simply lift R_p off ground and return it to an adjustable voltage V_X. By the superposition principle, we now have an apparent input error of $E_I + V_X$, and we can always adjust V_X to neutralize E_I. V_X is obtained from a dual reference source, such as the supply voltages if they are adequately regulated and filtered. In the circuits shown, we impose $R_B \gg R_C$ to avoid excessive loading at the wiper, and $R_A \ll R_p$ to avoid perturbing the existing resistance levels. The calibration procedure is similar to that for internal nulling.

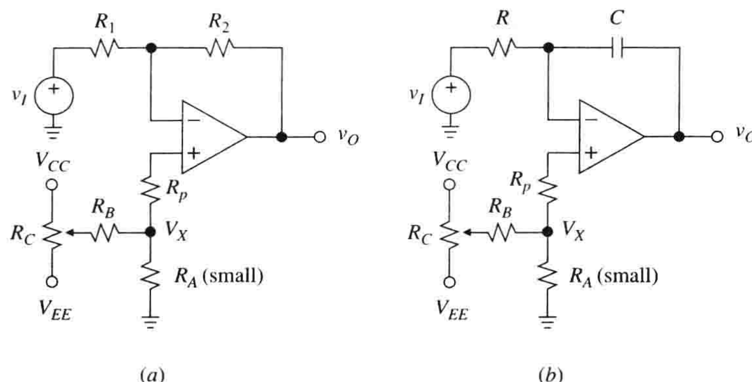


FIGURE 5.20
External offset-error nulling for (*a*) the inverting amplifier, and (*b*) the integrator.

EXAMPLE 5.8. A 741C op amp is to be used in the circuit of Fig. 5.20*a* to yield $A_s = -5$ V/V and $R_i = 30$ kΩ. Specify suitable resistances.

Solution. $R_1 = 30$ kΩ, $R_2 = 5R_1 = 150$ kΩ, and $R_p = R_1 \parallel R_2 = 25$ kΩ. Use the standard value $R_p = 24$ kΩ, and impose $R_A = 1$ kΩ to make up for the difference. We have $E_{I(\max)} = V_{OS(\max)} + (R_1 \parallel R_2)I_{OS(\max)} = 6\text{ mV} + (25\text{ k}\Omega) \times (200\text{ nA}) = 11$ mV. To be on the safe side, impose $-15\text{ mV} \le V_X \le 15$ mV. Thus, with the wiper all the way up, we want $R_A/(R_A + R_B) = (15\text{ mV})/(15\text{ V})$, or $R_B \cong 10^3 R_A = 1$ MΩ. Finally, choose $R_C = 100$ kΩ.

In principle, the foregoing scheme can be applied to any circuit that comes with a dc return to ground. In the circuit of Fig. 5.21, R_1 has been lifted off ground and returned to the adjustable voltage V_X. To avoid upsetting the signal gain, we must impose $R_{eq} \ll R_1$, where R_{eq} is the equivalent resistance of the nulling network as seen by R_1 (for $R_A \ll R_B$ we have $R_{eq} \cong R_A$.) Alternatively, we must decrease R_1 to the value $R_1 - R_{eq}$.

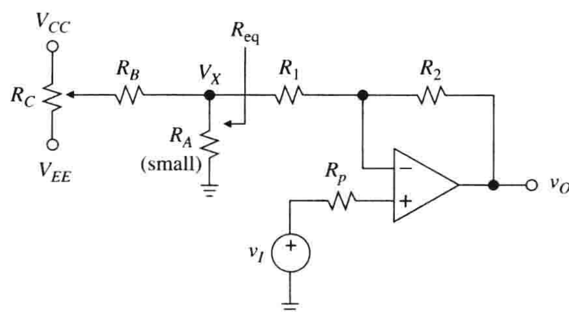


FIGURE 5.21
External offset-error nulling for the noninverting amplifier.

EXAMPLE 5.9. Assuming a 741C op amp in Fig. 5.21, specify suitable resistances for (*a*) $A_s = 5$ V/V, and (*b*) $A_s = 100$ V/V.

Solution.

(*a*) We want $A_s = 1 + R_2/R_1 = 5$, or $R_2 = 4R_1$. Pick $R_1 = 25.5$ kΩ, 1%, and $R_2 = 102$ kΩ, 1%. Then $R_p \cong 20$ kΩ. Moreover, $E_{O(\max)} = (1/\beta)E_{I(\max)} = 5[6 \text{ mV} + (20 \text{ k}\Omega) \times (200 \text{ nA})] = 50$ mV. To balance this out, we need $V_X = E_{O(\max)}/(-R_2/R_1) = 50/(-4) = -12.5$ mV. Pick a range of ± 15 mV to make sure. To avoid upsetting A_s, choose $R_A \ll R_1$, say, $R_A = 100$ Ω. Then, imposing $R_A/(R_A + R_B) = (15 \text{ mV})/(15 \text{ V})$ yields $R_B \cong 10^3 R_A = 100$ kΩ. Finally, let $R_C = 100$ kΩ.

(*b*) Now $1 + R_2/R_1 = 100$, or $R_2 = 99R_1$. Let $R_2 = 100$ kΩ, so $R_1 = 1010$ Ω. If we were to use $R_A = 100$ Ω as before, R_A would no longer be negligible compared to R_1. So let $R_1 = 909$ Ω, 1%, and $R_A = 1010 - 909 = 101$ Ω (use 102 Ω, 1%), so that the series $(R_1 + R_A)$ still ensures $A_s = 100$ V/V. Moreover, let $R_p \cong 1$ kΩ. Then, $E_{O(\max)} = 100[6 \text{ mV} + (1 \text{ k}\Omega) \times (200 \text{ nA})] = 620$ mV, and $V_X = E_{O(\max)}/(-R_2/R_1) = 620/(-10^5/909) = -5.6$ mV. Pick a range of ± 7.5 mV to make sure. Imposing $R_A/(R_A + R_B) = (7.5 \text{ mV})/(15 \text{ V})$ gives $R_B \cong 2000R_A \cong 200$ kΩ. Finally, let $R_C = 100$ kΩ.

In multiple-op amp circuits it is worth seeking ways of nulling the cumulative offset error with just one adjustment. A classic example is offered by the triple-op amp IA, where other critical parameters may also need adjustment, such as gain and CMRR.

In the circuit of Fig. 5.22, the voltage V_X is buffered by the low-output-impedance follower OA_4 to avoid upsetting bridge balance. The overall CMRR is the combined result of resistance mismatches and finite CMRRs of the individual op amps. At dc, where C_1 acts as an open circuit and R_9 has thus no effect, we adjust R_{10} to optimize

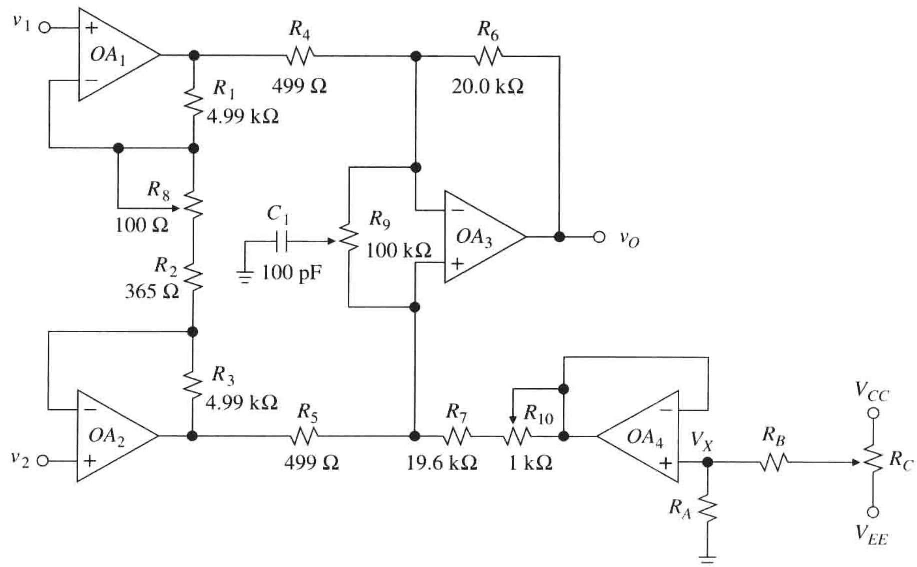


FIGURE 5.22
Instrumentation amplifier with $A = 1$ V/mV. (OA_1, OA_2, and OA_3: OP37C; OA_4: OP27; fixed resistances are 0.1%.)

the dc CMRR. At some high frequency, where C_1 provides a conductive path from R_9's wiper to ground, we adjust R_9 to deliberately unbalance the second stage and thus optimize the ac CMRR. The circuit is calibrated as follows:

1. With v_1 and v_2 grounded, adjust R_C for $v_O = 0$.
2. Adjust R_8 for the desired gain of 1000 V/V.
3. With the inputs tied together to a common source v_I, adjust R_{10} for the minimum change in v_O as v_I is switched from -10 V dc to $+10$ V dc.
4. With v_I a 10-kHz, 20-V peak-to-peak sine wave, adjust R_9 for the minimum ac component at the output.

EXAMPLE 5.10. Specify R_A, R_B, and R_C in Fig. 5.22, given the following maximum ratings for the OP37C low-noise precision high-speed op amp at $T = 25\ ^\circ\text{C}$: $I_B = 75$ nA, $I_{OS} = \pm 80$ nA, and $V_{OS} = 100\ \mu\text{V}$. Assume ±15-V supplies.

Solution. $E_{I1} = E_{I2} = V_{OS} + [R_1 \parallel (R_2 + R_8/2)]I_B = 10^{-4} + (5000 \parallel 208)75 \times 10^{-9} \cong 115\ \mu\text{V}$; $E_{I3} = 10^{-4} + (500 \parallel 20{,}000)80 \times 10^{-9} \cong 139\ \mu\text{V}$; $E_O = A(E_{I1} + E_{I2}) + (1/\beta_3)E_{I3} = 10^3 \times 2 \times 115 + (1 + 20/0.5)139 \cong 230\ \text{mV} + 5.7\ \text{mV} = 236\ \text{mV}$. According to Eq. (2.40) we need $-236\ \text{mV} \le V_X \le +236$ mV. Use 300 mV to make sure. Then, $R_A = 2\ \text{k}\Omega$, $R_B = 100\ \text{k}\Omega$, $R_C = 100\ \text{k}\Omega$.

Whether internal or external, nulling compensates only for the initial offset error V_{OS0}. As the operating conditions change, the error will reemerge, and if it rises above an intolerable level, it must be nulled periodically. The use of AZ or CS op amps may then be a preferable alternative.

5.7 INPUT VOLTAGE RANGE/OUTPUT VOLTAGE SWING

The *input voltage range* (IVR) of an op amp is the range of values of v_P and v_N over which the input stage will function properly, with all transistors operating in the active region, between the extremes of the edge of conduction and the edge of saturation. Similarly, the *output voltage swing* (OVS) is the range of values of v_O over which the output stage will function properly, with all transistors operating in the active region. Pushing v_O outside this range will cause it to saturate at V_{OH} or V_{OL}. So long as an op amp is operated within its IVR and OVS, and within its output current drive capability, it will force v_N to closely track v_P and give $v_O = a(v_P - v_N)$. Since the common-mode input voltage is $v_{IC} = (v_P - v_N)/2 \cong v_P$, the IVR is also called the *common-mode* IVR. Even though the IVR and OVS are given in the data sheets, a basic understanding of their origin will better help the user select a device for a given application. In the following discussion we investigate two representative cases, the bipolar 741 op amp and the two-stage CMOS op amp.

Input Voltage Range (IVR)

To find the IVR of the 741 op amp, refer to the input stage of Fig. 5.6, and use Fig. 5.23 (*a*) and (*b*) to visualize the subcircuits responsible for the IVR (note that R_1 has been ignored as it drops only about 10 mV). The upper limit of the IVR is the voltage

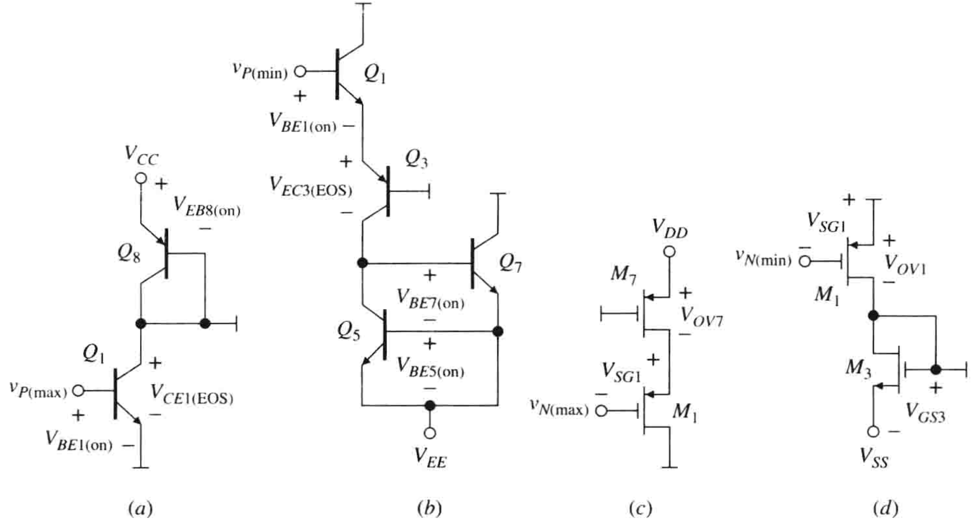


FIGURE 5.23
Subcircuits to find the IVR of the 741 op amp and of the two-stage CMOS op amp.

$v_{P(\max)}$ that brings Q_1 to the edge of saturation (EOS). Raising v_P further will drive Q_1 in saturation and subsequently turn off diode-connected Q_8, causing malfunction. By KVL, $v_{P(\max)} = V_{CC} - V_{EB8(\text{on})} - V_{CE1(\text{EOS})} + V_{BE1(\text{on})} \cong V_{CC} - V_{CE1(\text{sat})}$, where identical base-emitter voltage drops have been assumed. The lower limit is the voltage $v_{P(\min)}$ that brings Q_3 to the EOS. Lowering v_P further will drive Q_3 in saturation and subsequently turn off Q_1, causing malfunction. By KVL, $v_{P(\min)} = V_{EE} + V_{BE5(\text{on})} + V_{BE7(\text{on})} + V_{EC3(\text{EOS})} + V_{BE1(\text{on})} = V_{EE} + 3V_{BE(\text{on})} + V_{EC3(\text{EOS})}$. Summarizing,

$$v_{P(\max)} \cong V_{CC} - V_{CE1(\text{EOS})} \qquad v_{P(\min)} \cong V_{EE} + 3V_{BE(\text{on})} + V_{EC3(\text{EOS})} \tag{5.35}$$

Next, let us turn to the two-stage CMOS op amp of Fig. 5.4*a*, and use the subcircuits of Fig. 5.23 (*c*) and (*d*). Working with v_N, which in normal operation tracks v_P, we observe that the upper limit of the IVR is the voltage $v_{N(\max)}$ that brings M_7 to the EOS, where $V_{SD7} = V_{OV7}$, V_{OV7} being M_7's overdrive voltage. Raising v_N further will drive M_7 in the triode region and then turn off M_1, causing malfunction. The lower limit is the voltage $v_{N(\min)}$ that brings M_1 to the EOS where $V_{SD1} = V_{OV1}$. Lowering v_N further will drive M_1 in the triode region and then turn off diode-connected M_3, causing malfunction. KVL gives $v_{N(\max)} = V_{DD} - V_{OV7} - V_{SG1}$ and $v_{N(\min)} = V_{SS} + V_{GS3} + V_{OV1} - V_{SG1}$. But, by definition, $V_{SG1} = |V_{tp}| + V_{OV1}$, where V_{tp} is the threshold voltage of the *p*-MOSFETs, and $V_{GS3} = V_{tn} + V_{OV3}$, where V_{tn} is the threshold voltage of the *n*-MOSFETs. Consequently,

$$v_{N(\max)} = V_{DD} - V_{OV7} - |V_{tp}| - V_{OV1}$$
$$v_{N(\min)} = V_{SS} + V_{tn} + V_{OV3} - |V_{tp}| \tag{5.36}$$

EXAMPLE 5.11. (*a*) Find the IVR of the 741 op amp, assuming ±15-V supplies, 0.7-V base-emitter voltage drops, and 0.2-V EOS voltages. (*b*) Repeat, but for the CMOS op amp. Assume ±5-V supplies, 0.75-V threshold voltages, and 0.25-V overdrive voltages throughout.

Solution.

(*a*) Applying Eq. (5.35) we get $v_{P(\max)} \cong 15 - 0.2 = 14.8$ V, $v_{P(\min)} \cong -15 + 3 \times 0.7 + 0.2 = -12.7$ V, so the IVR can be expressed as $-12.7 \text{ V} \le v_{CM} \le 14.8$ V.

(*b*) Applying Eq. (5.36) we get $v_{N(\max)} = 5 - 0.25 - 0.75 - 0.25 = 3.75$ V, $v_{N(\min)} = -5 + 0.75 + 0.25 - 0.75 = -4.75$ V, so the IVR is $-4.75 \text{ V} \le v_{CM} \le 3.75$ V.

Output Voltage Swing (OVS)

The OVS of the 741 op amp is specified for a typical load $R_L = 2\ \text{k}\Omega$. With reference to the complete circuit schematic of Fig. 5A.2, we observe that v_O swings *positive* by the *pull-up* action by Q_{14}, as depicted in the subcircuit Fig. 5.24*a*, and v_O swings *negative* by the *pull-down* action by Q_{20}, as depicted in the subcircuit of Fig 5.24*b*. The upper limit of the OVS occurs when Q_{13} is driven to the EOS. Since $R_6 \ll R_L$, we approximate $v_{O(\max)} \cong V_{E14(\max)} = V_{CC} - V_{EC13(\text{EOS})} - V_{BE14(\text{on})}$. If Q_{13} is driven into full saturation, then v_O itself saturates at $V_{OH} \cong V_{CC} - V_{EC13(\text{sat})} - V_{BE14(\text{on})}$. The lower limit occurs when Q_{17} is driven to the EOS, so $v_{O(\min)} \cong V_{E20} = V_{EE} + V_{CE17(\text{EOS})} + V_{EB22(\text{on})} + V_{EB20(\text{on})}$. If Q_{17} is driven into full saturation, then v_O itself saturates at $V_{OL} \cong V_{EE} + V_{CE17(\text{sat})} + V_{EB22(\text{on})} + V_{EB20(\text{on})}$. Assuming the data of Example 5.11a, along with 0.1-V collector-emitter saturation voltages, the 741 has $v_{O(\max)} \cong 14.1$ V, $V_{OH} \cong 14.2$ V, $v_{O(\min)} \cong -13.4$ V, and $V_{OL} \cong -13.5$ V.

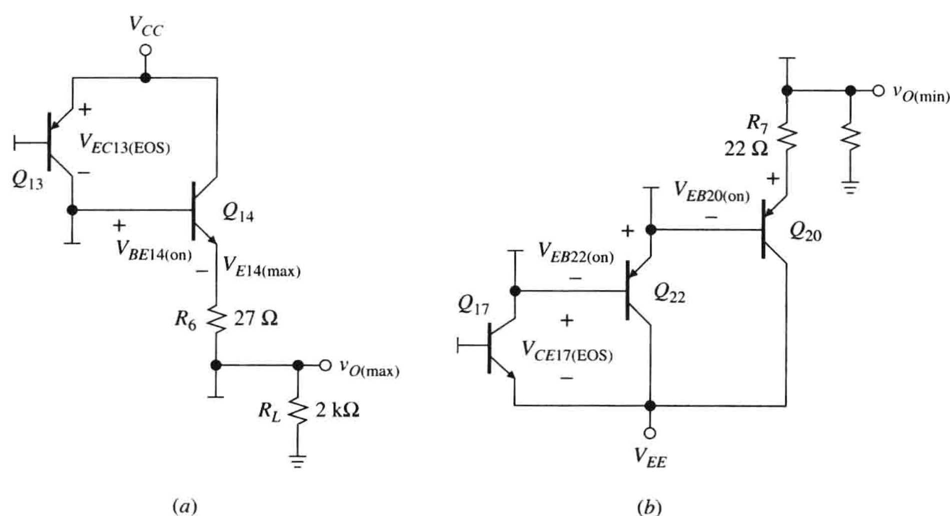


FIGURE 5.24
Subcircuits for the calculations of the OVS of the 741 op amp.

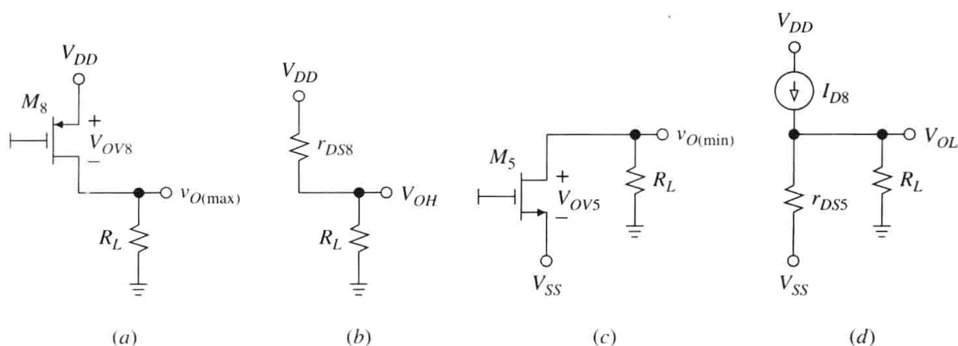


FIGURE 5.25
Subcircuits to find the OVS and saturation voltages of the two-stage CMOS op amp.

To find the OVS and saturation voltages of the two-stage CMOS op amp of Fig. 5.4*a*, refer to the subcircuits of Fig. 5.25. Positive swings of v_O occur under the pull-up action by M_8, so $v_{O(\max)} = V_{DD} - V_{OV8}$ (see Fig. 5.25*a*). Pushing v_O above $v_{O(\max)}$ will force M_8 into its triode region, where it acts as a resistor with a resistance[4] $r_{DS8} = 1/(k_8 V_{OV8})$. This resistance is fixed because $V_{OV8} = V_{OV6}$, which is also fixed. As shown in Fig. 5.25*b*, R_L forms a voltage divider with r_{DS8}, and only in the limit $R_L \to \infty$ will v_O swing all the way to the positive supply rail to give $V_{OH} \to V_{DD}$.

Negative swings of v_O occur under the pull-up action by M_5, so $v_{O(\min)} = V_{SS} + V_{OV5}$ (see Fig. 5.25*c*). Pushing v_O below $v_{O(\min)}$ will force M_5 into its triode region, leading to the situation of Fig. 5.25*d*. Note that in the limit $R_L \to \infty$ we have $V_{OL} \to V_{SS} + r_{DS5} I_{D5} \neq V_{SS}$, indicating this circuit's inability to swing v_O all the way to the negative supply rail.

Rail-to-Rail Op Amps

A combination of technological advances and innovative applications has resulted, over the years, in a progressive reduction of the power-supply voltages, especially in mixed-mode and portable systems (dual-supply or even single-supply voltages of just 1-2 V are becoming more and more common). In a low supply-voltage system it is imperative that the dynamic range of analog voltages be maximized, or, equivalently, that both the IVR and the OVR extend from *rail to rail* (or even a bit beyond the rails, if possible). The rail-to-rail concept is illustrated in Fig. 5.26.

Let us examine the IVR first. Figure 5.23 and Example 5.11 indicate that the IVR of the *npn* input pair extends almost all the way up to the *positive* supply, whereas that of a *p*-channel input pair extends almost all the way down to the *negative* supply. However, at the opposite supply side, the IVR is poor because of the headroom needed to keep the transistors in the forward-active region. In fact, we can generalize by stating that differential pairs made up of *npn* BJTs or *n*MOSFETs (call them *n*-type pairs) work well over the upper *portion* of the signal range, whereas their complementary pairs, made up of *pnp* BJTs or *p*MOSFETs (call them *p*-type pairs), work well over the *lower* portion. As we shall see shortly, an ingenious

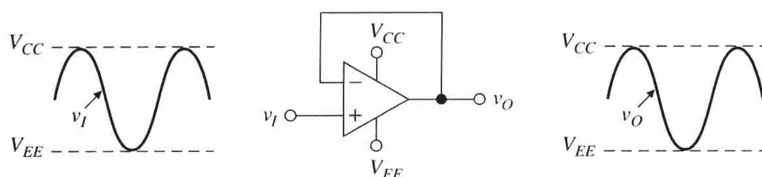


FIGURE 5.26
Waveforms for a voltage follower with rail-to-rail input and output capabilities.

solution is to utilize an n-type pair and a p-type pair in parallel to get the *best of both.*

Turning next to the OVS, we observe that both Fig. 5.24 and Fig. 5.25 refer to *push-pull* stages. However, the stage of Fig. 5.24, known as a common-source (CS) push-pull, offers better OVS characteristics than that of Fig. 5.23, in turn known as common-collector (CC) push-pull. This suggests operating the bipolar stage similarly, that is, as a common-emitter (CE) push-pull.

Figure 5.27 shows a bipolar implementation of the above concepts. When v_P and v_N are near V_{CC}, the p-pair Q_3-Q_4 is inactive, while the n-pair Q_1-Q_2 gives $v_{P(\max)} = V_{CC} - V_1 - V_{CE1(\text{EOS})} + V_{BE1(\text{on})}$. It is apparent that if the IC designer specifies the bias voltage V_{B56} so that $V_1 = V_{BE1(\text{on})} - V_{CE1(\text{EOS})} (\cong 0.7 - 0.2 = 0.5$ V), then $v_{P(\max)} = V_{CC}$. The dual situation occurs when v_P and v_N are near V_{EE}, in which case the n-pair Q_1-Q_2 is inactive, while the p-pair Q_3-Q_4

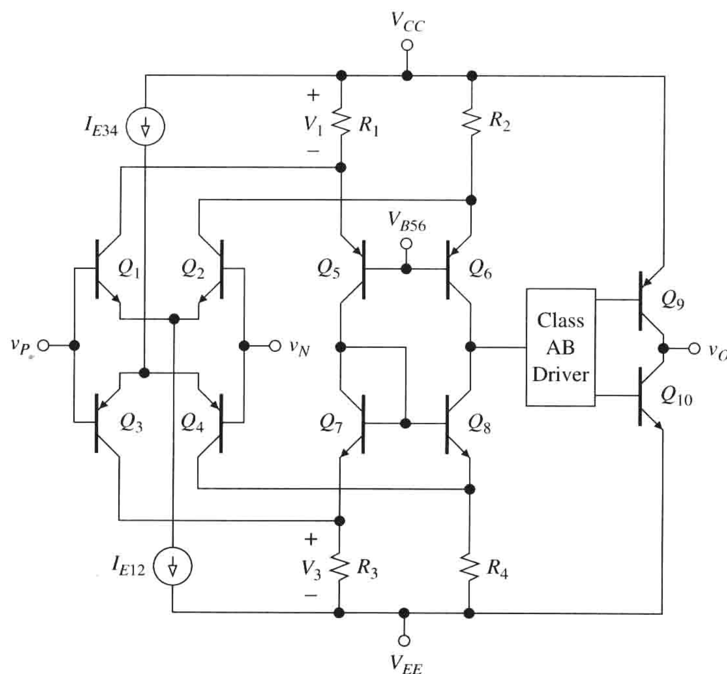


FIGURE 5.27
Simplified circuit diagram of a bipolar rail-to-rail op amp.

gives $v_{P(\min)} = V_{EE} + V_3 + V_{EC3(\text{EOS})} - V_{EB3(\text{on})}$. Again, if $V_3 = V_{EB3(\text{on})} - V_{EC3(\text{EOS})}$ ($\cong 0.7 - 0.2 = 0.5$ V), then $v_{P(\min)} = V_{EE}$, that is, we have a rail-to rail IVR!

With the CE push-pull output stage shown, the OVS limits are $v_{O(\max)} = V_{CC} - V_{EC9(\text{EOS})}$ and $v_{O(\min)} = V_{EE} + V_{CE10(\text{EOS})}$, typically within 0.2-V of the supply rails. Moreover, the saturation voltages are $V_{OH} = V_{CC} - V_{EC9(\text{sat})}$ and $V_{OL} = V_{EE} + V_{CE10(\text{sat})}$, typically within 0.1-V of the rails. The OVS superiority of the CC push-pull comes at the price of a high output resistance (compared with the very low output resistance of the CC push-pull). This renders the open-loop voltage gain strongly dependent on the output load.

In the CMOS version of Fig. 5.28, the bias voltages V_{G56} and V_{G78} are assumed to keep the M_{13}-M_{14} and M_9-M_{10} mirrors at the edge of saturation. Consequently, $v_{P(\max)} = V_{DD} - V_{OV13} - V_{OV1} + V_{GS1} = V_{DD} - V_{OV13} - V_{OV1} + (V_{tn} + V_{OV1}) = V_{DD} + V_{tn} - V_{OV13}$. To get an idea, if $V_{tn} = 0.75$ V and $V_{OV13} = 0.25$ V, then $v_{P(\max)} = V_{DD} + 0.5$ V, that is, the IVR extends a bit above the V_{DD} rail! Similarly, $v_{P(\min)} = V_{SS} + V_{OV9} - |V_{tp}|$. Assuming likewise $|V_{tp}| = 0.75$ V and $V_{OV9} = 0.25$ V gives $v_{P(\min)} = V_{SS} - 0.5$ V, a bit below the V_{SS} rail!

For the OVS and output saturation voltages, the considerations made in connection with Fig. 5.25 still hold, except that in true push-pull operation M_{11} is off at the lower end of the OVS, so in the limit $R_L \to \infty$ we have not only $V_{OH} \to V_{DD}$, but also $V_{OL} \to V_{SS}$, that is, true rail-to-rail performance!

The user needs to be aware of certain inherent limitations of the above circuits. As we sweep the common-mode input v_{CM} from the negative supply rail to the

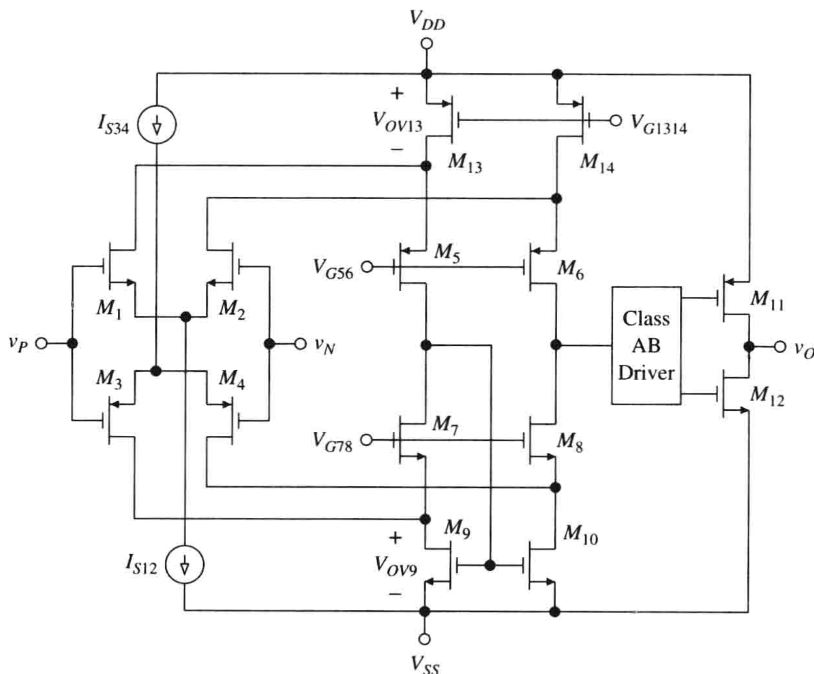


FIGURE 5.28
Simplified circuit diagram of a CMOS rail-to-rail op amp.

positive supply rail, the input stage goes through three modes of operation: near the low end the p-type pair is on and the n-type pair is off, in the middle region both pairs are simultaneously on, and near the high end the p-type pair is off and the n-type pair is on. Accordingly, we witness a change in the input offset voltage V_{OS}, a change in the input bias/offset currents I_B/I_{OS}, and a change in the overall input-stage g_m, which in turn affects the gain as well as the dynamics. For more details on these variations, along with clever optimization techniques, the reader is encouraged to consult the literature.[8]

5.8 MAXIMUM RATINGS

Like all electronic devices, op amps require that the user respect certain electrical and environmental limits. Exceeding these limits will generally result in malfunction or even damage. The range of operating temperatures over which op amp ratings are given are the *commercial range* (0 °C to + 70 °C), the *industrial range* (−25 °C to +85 °C), and the *military range* (−55 °C to +125 °C).

Absolute Maximum Ratings

These are the ratings that, if exceeded, are likely to cause permanent damage. The most important ones are the *maximum supply voltages,* the *maximum differential-mode* and *common-mode input voltages,* and the *maximum internal power dissipation* $P_{\max}$.

Figure 5A.1 indicates that for the 741C the maximum voltage ratings are, respectively, ±18 V, ±30 V, and ±15 V. (The large differential-mode rating of the 741 is made possible by the lateral *pnp* BJTs Q_3 and Q_4.) Exceeding these limits may trigger internal reverse-breakdown phenomena and other forms of electrical stress, whose consequences are usually detrimental, such as irreversible degradation of gain, input bias and offset currents, and noise, or permanent damage to the input stage. It is the user's responsibility to ensure that the device operates below its maximum ratings under all possible circuit and signal conditions.

Exceeding $P_{\max}$ will raise the chip temperature to intolerable levels and cause internal component damage. The value of $P_{\max}$ depends on the package type as well as the ambient temperature. The popular mini DIP package has $P_{\max} = 310$ mW up to 70 °C of ambient temperature, and derates linearly by 5.6 mW/°C beyond 70 °C.

EXAMPLE 5.12. What is the maximum current that a mini DIP 741C op amp is allowed to source at 0 V if $T \leq 70$ °C? If $T = 100$ °C?

Solution. From Fig. 5A.3 we find the supply current to be $I_Q = 2.8$ mA maximum. Recall from Section 1.8 that an op amp sourcing current dissipates $P = (V_{CC} - V_{EE})I_Q + (V_{CC} - V_O)I_O = 30 \times 2.8 + (15 - V_O)I_O$. Imposing $P \leq 310$ mV gives $I_O(V_O = 0) \leq (310 - 84)/15 \cong 15$ mA for $T \leq 70$ °C. For $T = 100$ °C we have $P_{\max} = 310 - (100 - 70)5.6 = 142$ mW, so now $I_O(V_O = 0) = (142 - 84)/15 \cong 3.9$ mA.

Overload Protection

To prevent excessive power dissipation in case of output overload, op amps are equipped with protective circuitry designed to limit the output current below a safety level called the *output short-circuit current* I_{sc}. The 741C has typically $I_{sc} \cong 25$ mA.

In the 741 diagram of Fig. 5A.2, overload protection is provided by the watchdog BJTs Q_{15} and Q_{21} and the current-sensing resistors R_6 and R_7. Under normal conditions these BJTs are off. However, should an output overload condition arise, such as an accidental short circuit, the resistance sensing the overload current will develop enough voltage to turn on the corresponding watchdog BJT; this, in turn, will limit the current through the corresponding output-stage BJT.

To illustrate with an example, suppose the op amp is designed to output a positive voltage, but an inadvertent output short forces v_O to 0 V, as depicted in Fig. 5.29. In response to this short, the second stage of the op amp will drive v_{B22} as positive as it can in a futile attempt to raise v_O. Consequently, Q_{22} will go off and let the entire bias current of 0.18 mA flow toward the base of Q_{14}. Were it not for the presence of Q_{15}, Q_{14} would amplify this current by β_{14} while sustaining $V_{CE} = V_{CC}$; the resulting power dissipation would most likely destroy it. However, with Q_{15} in place, only the current $i_{B14(\max)} = i_{C14(\max)}/\beta_{14} \cong [V_{BE15(\text{on})}/R_6]/\beta_{14}$ is allowed to reach the base of Q_{14}, the remainder being diverted to the output short via Q_{15}; hence, Q_{14} is protected.

With reference to Fig. 5A.2, we observe that just like Q_{15} protects Q_{14} when the op amp is *sourcing* current, Q_{21} protects Q_{20} during current *sinking*. However, since the base of Q_{20} is a low-impedance node because it is driven by emitter-follower Q_{22}, the action of Q_{21} is applied further upstream, via Q_{23}.

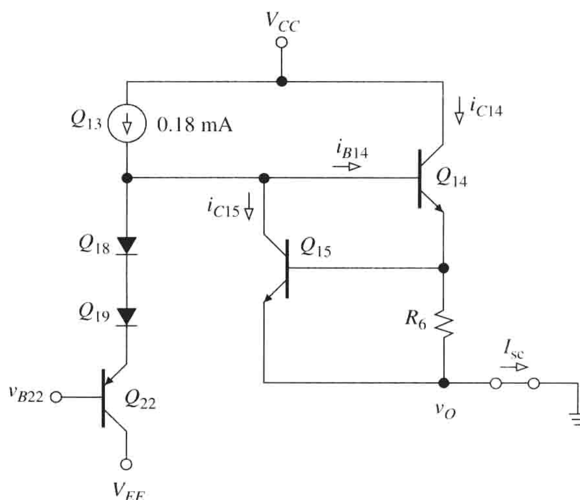


FIGURE 5.29
Partial illustration of overload protection circuitry for the 741 op amp.

EXAMPLE 5.13. Find all currents in the circuit of Fig. 5.29 if $R_6 = 27\ \Omega$, $\beta_{14} = \beta_{15} = 250$, and $V_{BE15(\text{on})} = 0.7$ V.

Solution. Q_{14} is limited to $I_{C14} = \alpha_{14} I_{E14} = \alpha_{14}[I_{R_6} + I_{B15}] \cong I_{R_6} = V_{BE15(\text{on})}/R_6 = 0.7/27 \cong 26$ mA. The current reaching the base of Q_{14} is $I_{B14} = I_{C14}/\beta_{14} = 26/250 \cong 0.104$ mA; the remainder, $I_{C15} = 0.18 - 0.104 \cong 76\ \mu$A, is diverted to the short. Hence, $I_{sc} \cong I_{C14} + I_{C15} \cong 26$ mA.

It is important to realize that during overload the actual output voltage is not what it should be: the protection circuitry prevents the op amp from properly influencing v_N, so during overload we generally have $v_N \neq v_P$.

Op amp types are available with much higher output current capabilities than the 741. Aptly referred to as *power op amps,* they are similar to their low-power counterparts except for the presence of heftier output stages and proper power packaging to handle the increased dissipation of heat. These op amps usually require heatsink mounting. Examples of power op amps are the PA04 and the OPA501, with peak output-current capabilities of 20 A and 10 A, respectively.

PROBLEMS

5.1 Simplified op amp circuit diagrams

5.1 Suppose the input stage of Fig. 5.2*a* has its input terminals and output terminal connected to ground. Moreover, let $I_A = 20\ \mu$A and $\beta_{Fp} = 50$, and let β_{Fn} be so large that we can ignore the base currents of Q_3 and Q_4. (*a*) Assuming perfectly matched transistor pairs, find the input-pin currents I_P and I_N and the output-pin current I_{O1}. (*b*) Repeat if $I_{s2} = 1.1 I_{s1}$ and $I_{s4} = I_{s3}$. (*c*) Find the voltage V_P that needs to be applied in order to drive I_{O1} to zero. What are the resulting values of I_P and I_N?

5.2 Input bias and offset currents

5.2 The circuit of Fig. 5.7*a* is to be used as an inverting amplifier with a gain of 10 V/V and is to employ the μA741C op amp. Specify suitable component values to ensure a maximum output error of 10 mV with minimum power dissipation in the resistors.

5.3 (*a*) Investigate the effect of I_B on the performance of the inverting amplifier of Fig. P1.64 if $I_B = 10$ nA and all resistances are 100 kΩ. (*b*) What dummy resistance R_p must be installed in series with the noninverting input to minimize E_O?

5.4 Investigate the effect of I_B and I_{OS} on the performance of the circuit of Fig. P1.17 if $I_B = 100$ nA flowing *out* of the op amp's input pins, and $I_{OS} = 10$ nA.

5.5 (*a*) Assuming the circuit of Fig. P1.65 has $R_1 = R_3 = R_5 = 10$ kΩ and $R_2 = R_4 = 20$ kΩ, find a dummy resistance that, inserted between the noninverting input and ground, will drive the output error E_O to zero for $I_{OS} = 0$. (*b*) What is E_O if $I_{OS} = 10$ nA?

5.6 The circuit of Fig. P5.6 exploits the matching properties of dual op amps to minimize the overall input current I_I. (*a*) Find the condition between R_2 and R_1 that yields $I_I = 0$ when the op amps are perfectly matched. (*b*) What if there is a 10% mismatch between the I_Bs of the op amps?

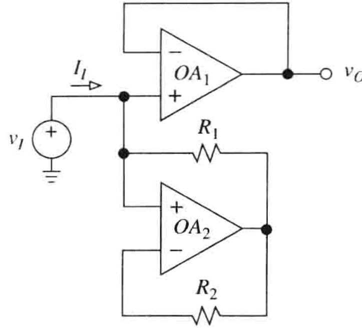


FIGURE P5.6

5.7 (*a*) Investigate the effect of I_{OS} on the performance of the Deboo integrator. (*b*) Assuming $C = 1$ nF and 100-kΩ resistances throughout, find $v_O(t)$ if $I_{OS} = \pm 1$ nA and $v_O(0) = 1$ V.

5.8 Investigate the effect of using an op amp with $I_B = 1$ nA and $I_{OS} = 0.1$ nA in the high-sensitivity *I-V* converter of Example 2.2. What dummy resistance R_p would you install in series with the noninverting input?

5.9 If $R_4/R_3 = R_2/R_1$, the circuit of Fig. P2.21 is a true *V-I* converter with $i_O = (R_2/R_1R_5) \times (v_2 - v_1)$ and $R_o = \infty$. What if the op amps have input bias currents I_{B1} and I_{B2}, and input offset currents I_{OS1} and I_{OS2}? Is i_O affected? Is R_o affected? How would you modify the circuit to optimize its dc performance?

5.10 Investigate the effect of I_B and I_{OS} in the current amplifier of Fig. 2.12*a*. How would you modify the circuit to minimize its dc error?

5.11 Assuming the multiple-feedback low-pass filter of Fig. 3.32 is in dc steady state (i.e., all transients have died out), investigate the effect of $I_B = 50$ nA if all resistances are 100 kΩ. What dummy resistor would you use to optimize the dc performance of the circuit? *Hint*: Set $V_i \to 0$.

5.12 (*a*) Assuming the low-pass filter of Example 3.9 is in dc steady state so all capacitors act as open circuits, find the output error E_O if $I_B = 50$ nA. *Hint*: Set $V_i \to 0$. (*b*) How would you modify the circuit so that $E_O = 0$ for $I_{OS} = 0$, without adding any extra circuit elements?

5.3 Low-input-bias-current op amps

5.13 A student is trying to figure out the technology of a certain unmarked op amp sample using the circuit of Fig. P5.13. Starting out with C completely discharged, the student monitors v_O with a digital voltmeter and finds that it takes 1 minute for v_O to rise from 0 V to 1 V. Moreover, after a sufficiently long time, v_O settles to 4 V. (*a*) What conclusions can the student draw? (*b*) How might the plot of the input bias current I_B over the range $-0.5 \text{ V} \le v_P \le 10.5$ V look like?

5.4 Input offset voltage

5.14 A FET-input op amp is connected as in Fig. 5.13*a* with $R_1 = 100$ Ω and $R_2 = 33$ kΩ, and gives $v_O = -0.5$ V. The same op amp is then moved to the circuit of Fig. 5.13*b* with

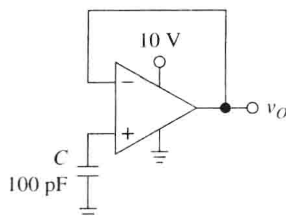


FIGURE P5.13

$R = 100\ \text{k}\Omega$ and $C = 1$ nF. Assuming $v_O(0) = 0$ and symmetric saturation voltages of ± 10 V, find the time it takes for the output to saturate.

5.15 If $R_4/R_3 = R_2/R_1$, the circuit of Fig. P2.22 is a true *V-I* converter with $i_O = R_2 v_I/(R_1 R_5)$ and $R_o = \infty$. What if the op amps have input offset voltages V_{OS1} and V_{OS2}, but are otherwise ideal? Is i_O affected? Is R_o affected?

5.16 In the circuit of Fig. 5.13*a* let $R_1 = 10\ \Omega$ and $R_2 = 100\ \text{k}\Omega$, and let the op amp have an offset drift of 5 μV/°C. (*a*) If the op amp has been trimmed for $v_O(25\ °\text{C}) = 0$, estimate $v_O(0\ °\text{C})$ and $v_O(70\ °\text{C})$. What do you expect their relative polarities to be? (*b*) If the same op amp is moved to the circuit of Fig. 5.13*b* with $R = 100\ \text{k}\Omega$ and $C = 1$ nF, find $v_O(t)$ both at 0 °C and at 70 °C.

5.17 Investigate the effect of using an op amp with $\text{CMRR}_{\text{dB}} = 100$ dB on the output resistance of a Howland current pump made up of four perfectly matched 10-kΩ resistances. Except for CMRR, the op amp is ideal.

5.18 Investigate the effect of using an op amp with $V_{OS0} = 100\ \mu$V and $\text{CMRR}_{\text{dB}} = 100$ dB in a Deboo integrator that uses four perfectly matched 100-kΩ resistances and a 1-nF capacitance (assume $a = \infty$). *Hint:* Find the Norton equivalent seen by the capacitance for $v_I = 0$.

5.19 Suppose the difference amplifier of Fig. 2.13*a* is implemented with a JFET op amp and two perfectly matched resistor pairs $R_3 = R_1 = 1.0\ \text{k}\Omega$ and $R_4 = R_2 = 100.0\ \text{k}\Omega$. The inputs v_1 and v_2 are tied together and driven by a common voltage v_{IC}. (*a*) If it is found that with $v_{IC} = 0$ V the circuit gives $v_O = 5.0$ mV, and with $v_{IC} = 2.0$ V it gives $v_O = -1.0$ mV, what is the CMRR? (*b*) If it is found that lowering the supply voltages by 0.5 V with v_{IC} still at 2.0 V results in $v_O = +1.0$ mV, what is the PSRR?

5.20 Suppose the *I-V* converter of Fig. 2.2 is implemented with $R = 500\ \text{k}\Omega$, $R_1 = 1.0\ \text{k}\Omega$, and $R_2 = 99\ \text{k}\Omega$, and a FET-input op amp. The op amp has $a = 100$ dB, PSRR = 80 dB, and $V_{OS0} = 0$, and is powered from dual supplies $\pm V_S = \pm[5 + 1\sin(2\pi t)]$ V. (*a*) Sketch and label the circuit, and following the reasoning in connection with Fig. 5.14, convert it to an equivalent circuit that uses an op amp with $a = \infty$ but with a suitable input offset voltage V_{OS} to account for the finite a and finite PSRR. (*b*) If we write $v_O = A_{\text{ideal}} i_I + E_O$, what are A_{ideal} and E_O?

5.21 Assuming perfectly matched resistances in the difference amplifier of Fig. 2.13*a*, show that if we define the CMRR of the op amp as $1/\text{CMRR}_{\text{OA}} = \partial V_{OS}/\partial v_{\text{CM(OA)}}$ and that of the difference amplifier as $1/\text{CMRR}_{\text{DA}} = A_{\text{cm}}/A_{\text{dm}}$, where $A_{\text{cm}} = \partial v_O/\partial v_{\text{CM(DA)}}$ and $A_{\text{dm}} = R_2/R_1$, we then have $\text{CMRR}_{\text{DA}} = \text{CMRR}_{\text{OA}}$.

5.22 The difference amplifier of Problem 5.21 uses a 741 op amp with $R_1 = 1\ \text{k}\Omega$ and $R_2 = 100\ \text{k}\Omega$. Find the worst-case CMRR of the circuit for the case of (*a*) perfectly matched resistances, and (*b*) 1% resistances. Comment.

5.23 In the difference amplifier of Problem 5.22 the inputs are tied together and are driven by $v_{CM} = 1 \sin(2\pi ft)$ V. Using the CMRR plot of Fig. 5A.6, predict the output at $f =$ 1 Hz, 1 kHz, and 10 kHz.

5.24 (*a*) Assuming perfectly matched op amps and resistances in the dual-op-amp IA of Fig. 2.23, show that if we define the CMRR of each op amp as $1/\text{CMRR}_{\text{OA}} = \partial V_{OS}/\partial v_{\text{CM(OA)}}$ and that of the IA as $1/\text{CMRR}_{\text{IA}} = A_{\text{cm}}/A_{\text{dm}}$, where $A_{\text{cm}} = \partial v_O/\partial v_{\text{CM(DA)}}$ and $A_{\text{dm}} = 1 + R_2/R_1$, then we have $\text{CMRR}_{\text{IA(min)}} = 0.5 \times \text{CMRR}_{\text{OA(min)}}$. (*b*) If an IA with a gain of 100 V/V is implemented with perfectly matched resistances and a dual OP227A op amp ($\text{CMRR}_{\text{dB}} = 126$ dB typical, 114 dB minimum), find the worst-case output change for a 10-V common-mode input change. What is the corresponding A_{cm}?

5.25 Assuming perfectly matched op amps and resistances in the triple-op amp IA of Fig. 2.20, derive a relationship between $\text{CMRR}_{\text{IA(min)}}$ and $\text{CMRR}_{\text{OA(min)}}$, where $1/\text{CMRR}_{\text{OA}} = \partial V_{OS}/\partial v_{\text{CM(OA)}}$, and $1/\text{CMRR}_{\text{IA}} = A_{\text{cm}}/A_{\text{dm}}$.

5.26 In the inverting integrator of Fig. 1.20 let $R = 100\ \text{k}\Omega$, $C = 10$ nF, and $v_I = 0$, and let the capacitor be initially charged such that $v_O(t = 0) = 10$ V. Except for a finite open-loop gain of 10^5 V/V, the op amp is ideal. Find $v_O(t > 0)$.

5.27 An op amp with $a_{\min} = 10^4$ V/V, $V_{OS0(\max)} = 2$ mV, and $\text{CMRR}_{\text{dB(min)}} = \text{PSRR}_{\text{dB(min)}} =$ 74 dB is configured as a voltage follower. (*a*) Estimate the worst-case departure of v_O from the ideal for $v_I = 0$ V. (*b*) Repeat with $v_I = 10$ V. (*c*) Repeat if the supply voltages are lowered from ± 15 V to ± 12 V.

5.5 Low-input-offset-voltage op amps

5.28 (*a*) With reference to the input stage of Fig. 5.2*a*, show that in case of mismatches in the Q_1-Q_2 and Q_3-Q_4 pairs, the voltage V_{OS} needed to drive i_{O1} to zero is

$$V_{OS} = V_T \ln\left(\frac{I_{s1}}{I_{s2}}\frac{I_{s4}}{I_{s3}}\right)$$

Hint: Use $v_N = v_P + V_{OS}$. (*b*) Assuming $V_T = 26$ mV, calculate V_{OS} for the following 10% mismatch cases: $(I_{s1}/I_{s2})(I_{s4}/I_{s3}) = (1.1)(1.1)$, $(1/1.1)(1.1)$, $(1.1)(1/1.1)$, $(1/1.1)(1/1.1)$. Compare with Eq. (5.28) and comment.

5.29 Suppose both the BJTs and the corresponding collector resistance strings in the circuit of Fig. 5.16 are mismatched. If $I_A = 100\ \mu$A, and Q_1's emitter area is 7.5% larger than Q_2's, and if it is found that $V_{C1} - V_{C2} = -15$ mV, which resistor string needs to be trimmed to drive the difference $V_{C1} - V_{C2}$ to zero, and by how many ohms?

5.6 Input offset error and compensation techniques

5.30 Consider the circuit obtained from that of Fig. 1.42*a* by letting $v_I = 0$ and replacing R_3 with a 10-nF capacitor. (*a*) If $R_1 = R_2 = R_4 = 10\ \text{k}\Omega$, and the op amp has $V_{OS} = 1.0$ mV and $I_B = 50$ nA flowing into the device, obtain an expression for $v_O(t)$. (*b*) Assuming the capacitor is initially discharged, and the op amp saturates at

± 10 V, find the time it takes for the op amp to reach saturation. (*c*) Assuming the op amp allows for internal offset trimming, how would you proceed to compensate it?

5.31 Repeat Example 5.8, but for the integrator of Fig. 5.20*b* for the case $R = 100\ \text{k}\Omega$.

5.32 In the noninverting amplifier of Fig. 1.15*a* let $R_1 = 10\ \Omega$, $R_2 = 10\ \text{k}\Omega$, and $v_I = 0$. The output v_O is monitored with a voltmeter and is found to be 0.480 V. If inserting a 1-MΩ resistor in series with the noninverting input pin gives $v_O = 0.780$ V, but inserting it in series with the inverting input pin gives $v_O = 0.230$ V, find I_B, I_{OS}, and V_{OS}. What is the direction of I_B?

5.33 Figure P5.33 shows a widely used test fixture to characterize the op amp referred to as *device under test* (DUT). The purpose of OA_2, which is assumed ideal, is to keep DUT's output near 0 V, or in the middle of the linear region. Find V_{OS0}, I_P, I_N, I_B, I_{OS}, and the gain a for the DUT, given the following measurements: (*a*) $v_2 = -0.75$ V with SW_1 and SW_2 closed and $v_1 = 0$ V; (*b*) $v_2 = +0.30$ V with SW_1 closed, SW_2 open, and $v_1 = 0$ V; (*c*) $v_2 = -1.70$ V with SW_1 open, SW_2 closed, and $v_1 = 0$ V; (*d*) $v_2 = -0.25$ V with SW_1 and SW_2 closed, and $v_1 = -10$ V.

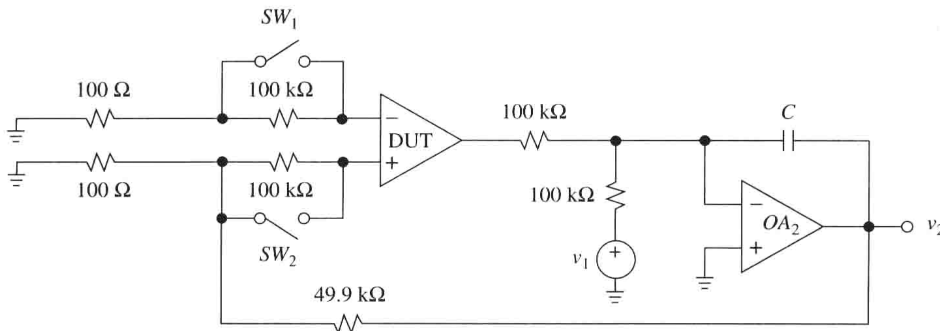


FIGURE P5.33

5.34 (*a*) In the circuit of Fig. P1.15 obtain an expression for the output error E_O as a function of I_P, I_N, and V_{OS}. (*b*) Repeat, but for the circuit of Fig. P1.16. *Hint:* In each case set the independent source to zero.

5.35 Repeat Problem 5.34, but for the circuits of Figs. P1.18 and P1.19.

5.36 (*a*) Find output error E_O for the *I-V* converter of Fig. 2.1. (*b*) Repeat if the noninverting input pin is returned to ground via a dummy resistance $R_p = R$. (*c*) Devise a scheme for the external nulling of E_O if $R = 1\ \text{M}\Omega$, $I_{OS} = 1$ nA maximum, and $V_{OS} = 1$ mV maximum.

5.37 What input-stage technology would you choose for the op amp of the high-sensitivity *I-V* converter of Example 2.2? How would you modify the circuit for a minimum output error E_O? How would you make provision for the external nulling of E_O?

5.38 Using the OP227A dual-precision op amp ($V_{OS(\max)} = 80\ \mu$V, $I_{B(\max)} = \pm 40$ nA, $I_{OS(\max)} = 35$ nA, and $\text{CMRR}_{\text{dB}(\min)} = 114$ dB), design a dual-op amp IA with a gain of 100 V/V. Assuming perfectly matched resistances, what is the maximum output error for $v_1 = v_2 = 0$? For $v_1 = v_2 = 10$ V?

5.39 If $R_2 + R_3 = R_1$, the circuit of Fig. P2.33 is a true *V-I* converter with $i_O = v_I/R_3$ and $R_o = \infty$. What if the op amps have nonzero input bias and offset currents and offset voltages? Is i_O affected? Is R_o affected? How would you make provisions for minimizing the total error? For externally nulling it?

5.40 (*a*) Investigate the effect of the offset voltages V_{OS1} and V_{OS2} on the performance of the dual-op-amp transducer amplifier of Fig. 2.40 for the case $\delta = 0$. (*b*) Devise a scheme to externally null the output offset error, and illustrate how it works.

5.41 Repeat Problem 5.40, but for the transducer amplifier of Fig. P2.62.

5.42 An *I-V* converter with a sensitivity of 1 V/μA is to be designed using an op amp with $V_{OS(\max)} = 1$ mV and $I_{OS(\max)} = 2$ nA. Two alternatives are being evaluated, namely, the circuit of Fig. 2.1 with $R = 1$ MΩ, and the circuit of Fig. 2.2 with $R = 100$ kΩ, $R_1 = 2.26$ kΩ, and $R_2 = 20$ kΩ; both circuits use an appropriate dummy resistance R_p to minimize the error due to I_B. Which circuit is preferable from the viewpoint of minimizing the untrimmed output error? What is the main reason?

5.43 Assuming the multiple-feedback band-pass filter of Example 3.15 is in dc steady state (i.e., all transients have died out), investigate the effect of $I_B = 50$ nA, $I_{OS} = 5$ nA, and $V_{OS} = 1$ mV upon the circuit's performance. How would you modify the circuit to minimize the output error? To null it? *Hint:* Assume a zero input.

5.44 Repeat Problem 5.43, but for the low-pass *KRC* filter of Example 3.8.

5.45 Repeat Problem 5.43, but for the band-bass and band-reject *KRC* filters of Examples 3.13 and 3.14.

5.46 The biquad filter of Example 3.19 is implemented with FET-input op amps having maximum input offset voltages of 5 mV. Investigate the effect on circuit performance and devise a method to trim the output dc error for the low-pass output.

5.7 Input voltage range/output voltage swing

5.47 Find the IVR and OVS of the folded-cascode CMOS op amp of Fig. 5.4*b*. Assume ±5-V supplies, 0.75-V threshold voltages, and 0.25-V overdrive voltages throughout. Also, assume the I_{SS} source is generated by a current mirror similar to the M_6-M_7 mirror of Fig. 5.4*a*, and V_{BIAS} is such that the I_{BIAS} current sinks drop 0.25-V each.

5.48 Suppose OA_2 in Fig. 1.47 is a true rail-to-rail op amp configured for a closed-loop gain of −2 V/V. (*a*) Sketch and label v_I, v_O, and v_D if $v_I = (1.0\text{ V})\sin(2\pi 10^3 t)$. (*b*) Repeat if $v_I = (1.5\text{ V})\sin(2\pi 10^3 t)$. (*c*) What is the maximum sinusoidal input that will still result in an undistorted output?

5.8 Maximum ratings

5.49 A 741 op amp is connected as a voltage follower and programmed to give $v_O = 10$ V. Using the simplified circuit of Fig. 5.29 with $R_6 = 27\ \Omega$, β_{FS} of 250, and base-emitter junction drops of 0.7 V, find v_{B22}, i_{C14}, i_{C15}, $P_{Q_{14}}$, and v_O if the output load is (*a*) $R_L = 2$ kΩ, and (*b*) $R_L = 200\ \Omega$.

REFERENCES

1. J. E. Solomon, "The Monolithic Operational Amplifier: A Tutorial Study," *IEEE J. Solid-State Circuits,* Vol. SC-9, December 1974, pp. 314–332.
2. W. Jung, *Op Amp Applications Handbook* (Analog Devices Series), Elsevier/Newnes, Oxford, UK, 2005.
3. P. R. Gray, P. J. Hurst, S. H. Lewis, and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 5th ed., John Wiley & Sons, New York, 2009.
4. S. Franco, *Analog Circuit Design—Discrete and Integrated,* McGraw-Hill, New York, 2014.
5. G. R. Boyle, B. M. Cohn, D. O. Pederson, and J. E. Solomon, "Macromodeling of Integrated Circuit Operational Amplifiers," *IEEE J. Solid-State Circuits,* Vol. SC-9, December 1974, pp. 353–363.
6. C. C. Enz and G. C. Temes, "Circuit Techniques for Reducing the Effects of Op-Amp Imperfections: Autozeroing, Correlated Double Sampling, and Chopper Stabilization," *IEEE Proceedings,* Vol. 84, No. 11, November 1996, pp. 1584–1614.
7. J. Williams, "Chopper-Stabilized Monolithic Op Amp Suits Diverse Uses," *EDN,* Feb. 21, 1985, pp. 305–312; and "Chopper Amplifier Improves Operation of Diverse Circuits," *EDN,* Mar. 7, 1985, pp. 189–207.
8. J. Huijsing, *Operational Amplifiers—Theory and Design,* 2nd ed., Springer, Dordrecht, 2011.

APPENDIX 5A
DATA SHEETS OF THE μA741 OP AMP*



FAIRCHILD
A Schlumberger Company

μA741
Operational Amplifier

Linear Division Operational Amplifiers

Description
The μA741 is a high performance monolithic operational amplifier constructed using the Fairchild Planar Epitaxial process. It is intended for a wide range of analog applications. High common mode voltage range and absence of latch up tendencies make the μA741 ideal for use as a voltage follower. The high gain and wide range of operating voltage provide superior performance in integrator, summing amplifier, and general feedback applications.

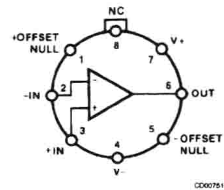
- **No Frequency Compensation Required**
- **Short Circuit Protection**
- **Offset Voltage Null Capability**
- **Large Common Mode And Differential Voltage Ranges**
- **Low Power Consumption**
- **No Latch Up**

Absolute Maximum Ratings

Storage Temperature Range	
Metal Can and Ceramic DIP	−65°C to +175°C
Molded DIP and SO-8	−65°C to +150°C
Operating Temperature Range	
Extended (μA741AM, μA741M)	−55°C to +125°C
Commercial (μA741EC, μA741C)	0°C to +70°C
Lead Temperature	
Metal Can and Ceramic DIP (soldering, 60 s)	300°C
Molded DIP and SO-8 (soldering, 10 s)	265°C
Internal Power Dissipation[1, 2]	
8L-Metal Can	1.00 W
8L-Molded DIP	0.93 W
8L-Ceramic DIP	1.30 W
SO-8	0.81 W
Supply Voltage	
μA741A, μA741, μA741E	±22 V
μA741C	±18 V
Differential Input Voltage	±30 V
Input Voltage[3]	±15 V
Output Short Circuit Duration[4]	Indefinite

Notes
1. $T_{J\,Max}$ = 150°C for the Molded DIP and SO-8, and 175°C for the Metal Can and Ceramic DIP.
2. Ratings apply to ambient temperature at 25°C. Above this temperature, derate the 8L-Metal Can at 6.7 mW/°C, the 8L-Molded DIP at 7.5 mW/°C, the 8L-Ceramic DIP at 8.7 mW/°C, and the SO-8 at 6.5 mW/°C.
3. For supply voltages less than ±15 V, the absolute maximum input voltage is equal to the supply voltage.
4. Short circuit may be to ground or either supply. Rating applies to 125°C case temperature or 75°C ambient temperature.

Connection Diagram
8-Lead Metal Package
(Top View)

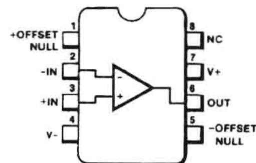


Lead 4 connected to case.

Order Information

Device Code	Package Code	Package Description
μA741HM	5W	Metal
μA741HC	5W	Metal
μA741AHM	5W	Metal
μA741EHC	5W	Metal

Connection Diagram
8-Lead DIP and SO-8 Package
(Top View)



Order Information

Device Code	Package Code	Package Description
μA741RM	6T	Ceramic DIP
μA741RC	6T	Ceramic DIP
μA741SC	KC	Molded Surface Mount
μA741TC	9T	Molded DIP
μA741ARM	6T	Ceramic DIP
μA741ERC	6T	Ceramic DIP
μA741ETC	9T	Molded DIP

FIGURE 5A.1

μA741

Equivalent Circuit

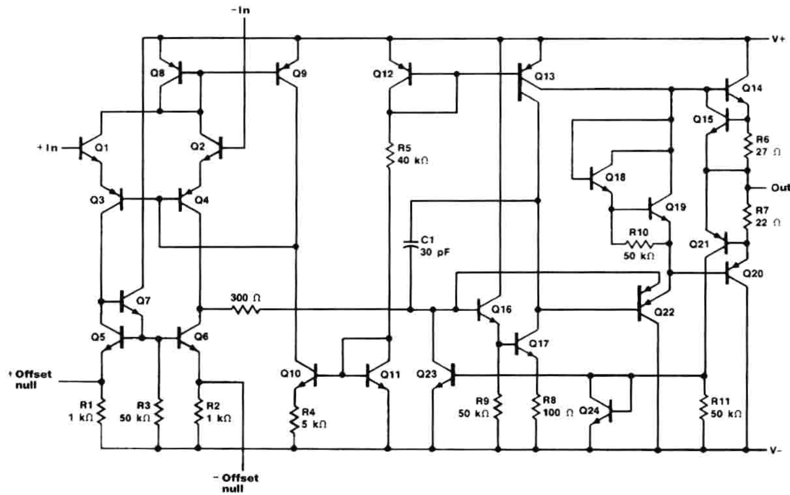


FIGURE 5A.2

μA741

μA741 and μA741C
Electrical Characteristics $T_A = 25°C$, $V_{CC} = \pm 15$ V, unless otherwise specified.

Symbol	Characteristic	Condition	μA741 Min	μA741 Typ	μA741 Max	μA741C Min	μA741C Typ	μA741C Max	Unit
V_{IO}	Input Offset Voltage	$R_S \leq 10$ kΩ		1.0	5.0		2.0	6.0	mV
$V_{IO\ adj}$	Input Offset Voltage Adjustment Range			±15			±15		mV
I_{IO}	Input Offset Current			20	200		20	200	nA
I_{IB}	Input Bias Current			80	500		80	500	nA
Z_I	Input Impedance		0.3	2.0		0.3	2.0		MΩ
I_{CC}	Supply Current			1.7	2.8		1.7	2.8	mA
P_c	Power Consumption			50	85		50	85	mW
CMR	Common Mode Rejection		70			70	90		dB
V_{IR}	Input Voltage Range		±12	±13		±12	±13		V
PSRR	Power Supply Rejection Ratio			30	150				μV/V
		$V_{CC} = \pm 5.0$ V to ± 18 V					30	150	
I_{OS}	Output Short Circuit Current			25			25		mA
A_{VS}	Large Signal Voltage Gain	$R_L \geq 2.0$ kΩ, $V_O = \pm 10$ V	50	200		20	200		V/mV
V_{OP}	Output Voltage Swing	$R_L = 10$ kΩ	±12			±12	±14		V
		$R_L = 2.0$ kΩ	±10			±10	±13		
TR	Transient Response: Rise time	$V_I = 20$ mV, $R_L = 2.0$ kΩ, $C_L = 100$ pF, $A_V = 1.0$		0.3			0.3		μs
	Transient Response: Overshoot			5.0			5.0		%
BW	Bandwidth			1.0			1.0		MHz
SR	Slew Rate	$R_L \geq 2.0$ kΩ, $A_V = 1.0$		0.5			0.5		V/μs

FIGURE 5A.3

μA741 and μA741C (Cont.)
Electrical Characteristics Over the range of $-55°C \leqslant T_A \leqslant +125°C$ for μA741, $0°C \leqslant T_A \leqslant +70°C$ for μA741C, unless otherwise specified.

Symbol	Characteristic	Condition	μA741			μA741C			Unit
			Min	Typ	Max	Min	Typ	Max	
V_{IO}	Input Offset Voltage							7.5	mV
		$R_S \leqslant 10$ kΩ		1.0	6.0				
$V_{IO\ adj}$	Input Offset Voltage Adjustment Range			±15			±15		mV
I_{IO}	Input Offset Current							300	nA
		$T_A = +125°C$		7.0	200				
		$T_A = -55°C$		85	500				
I_{IB}	Input Bias Current							800	nA
		$T_A = +125°C$		0.03	0.5				μA
		$T_A = -55°C$		0.3	1.5				
I_{CC}	Supply Current	$T_A = +125°C$		1.5	2.5				mA
		$T_A = -55°C$		2.0	3.3				
P_c	Power Consumption	$T_A = +125°C$		45	75				mW
		$T_A = -55°C$		60	100				
CMR	Common Mode Rejection	$R_S \leqslant 10$ kΩ	70	90					dB
V_{IR}	Input Voltage Range		±12	±13					V
PSRR	Power Supply Rejection Ratio			30	150				μV/V
A_{VS}	Large Signal Voltage Gain	$R_L \geqslant 2.0$ kΩ, $V_O = \pm 10$ V	25			15			V/mV
V_{OP}	Output Voltage Swing	$R_L = 10$ kΩ	±12	±14					V
		$R_L = 2.0$ kΩ	±10	±13		±10	±13		

FIGURE 5A.4

μA741

μA741A and μA741E
Electrical Characteristics $T_A = 25°C$, $V_{CC} = \pm 15$ V, unless otherwise specified.

Symbol	Characteristic		Condition	Min	Typ	Max	Unit
V_{IO}	Input Offset Voltage		$R_S \leq 50\ \Omega$		0.8	3.0	mV
I_{IO}	Input Offset Current				3.0	30	nA
I_{IB}	Input Bias Current				30	80	nA
Z_I	Input Impedance		$V_{CC} = \pm 20$ V	1.0	6.0		MΩ
P_c	Power Consumption		$V_{CC} = \pm 20$ V		80	150	mW
PSRR	Power Supply Rejection Ratio		$V_{CC} = +10$ V, -20 V to $V_{CC} = +20$ V, -10 V, $R_S = 50\ \Omega$		15	50	μV/V
I_{OS}	Output Short Circuit Current			10	25	40	mA
A_{VS}	Large Signal Voltage Gain		$V_{CC} = \pm 20$ V, $R_L \geq 2.0$ kΩ, $V_O = \pm 15$ V	50	200		V/mV
TR	Transient Response	Rise time	$A_V = 1.0$, $V_{CC} = \pm 20$ V, $V_I = 50$ mV, $R_L = 2.0$ kΩ, $C_L = 100$ pF		0.25	0.8	μs
		Overshoot			6.0	20	%
BW	Bandwidth			0.437	1.5		MHz
SR	Slew Rate		$V_I = \pm 10$ V, $A_V = 1.0$	0.3	0.7		V/μs

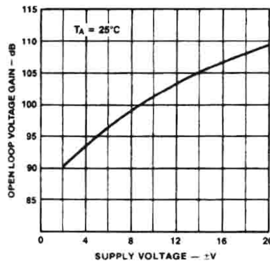
The following specifications apply over the range of $-55°C \leq T_A \leq +125°C$ for the μA741A, and $0°C \leq T_A \leq +70°C$ for the μA741E.

Symbol	Characteristic	Condition			Min	Typ	Max	Unit
V_{IO}	Input Offset Voltage						4.0	mV
$\Delta V_{IO}/\Delta T$	Input Offset Voltage Temperature Sensitivity						15	μV/°C
$V_{IO\ adj}$	Input Offset Voltage Adjustment Range	$V_{CC} = \pm 20$ V			10			mV
I_{IO}	Input Offset Current						70	nA
$\Delta I_{IO}/\Delta T$	Input Offset Current Temperature Sensitivity						0.5	nA/°C
I_{IB}	Input Bias Current						210	nA
Z_I	Input Impedance				0.5			MΩ
P_c	Power Consumption	$V_{CC} = \pm 20$ V	μA741A	−55°C			165	mW
				+125°C			135	
			μA741E				150	
CMR	Common Mode Rejection	$V_{CC} = \pm 20$ V, $V_I = \pm 15$ V, $R_S = 50\ \Omega$			80	95		dB
I_{OS}	Output Short Circuit Current				10		40	mA
A_{VS}	Large Signal Voltage Gain	$V_{CC} = \pm 20$ V, $R_L \geq 2.0$ kΩ, $V_O = \pm 15$ V			32			V/mV
		$V_{CC} = \pm 5.0$ V, $R_L \geq 2.0$ kΩ, $V_O = \pm 2.0$ V			10			
V_{OP}	Output Voltage Swing	$V_{CC} = \pm 20$ V	$R_L = 10$ kΩ		±16			V
			$R_L = 2.0$ kΩ		±15			

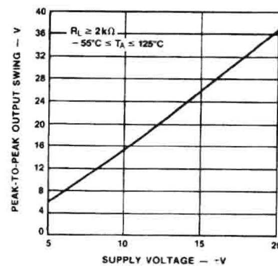
FIGURE 5A.5

Typical Performance Curves

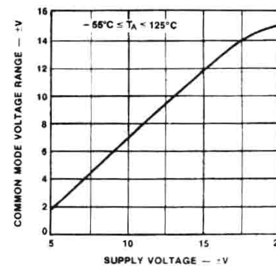
Voltage Gain vs Supply Voltage for μA741/A



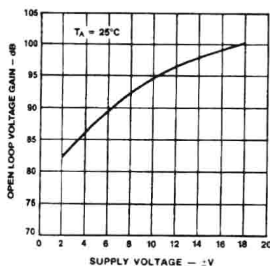
Output Voltage Swing vs Supply Voltage for μA741/A



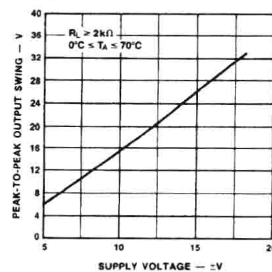
Input Common Mode Voltage vs Supply Voltage for μA741/A



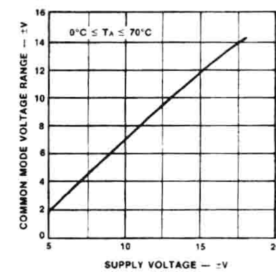
Voltage Gain vs Supply Voltage for μA741C/E



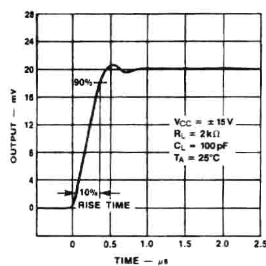
Output Voltage Swing vs Supply Voltage for μA741C/E



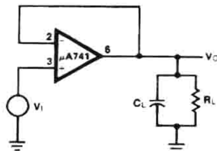
Input Common Mode Voltage Range vs Supply Voltage for μA741C/E



Transient Response for μA741C/E



Transient Response Test Circuit for μA741C/E



Lead numbers are shown for metal package only

Common Mode Rejection Ratio vs Frequency for μA741C/E

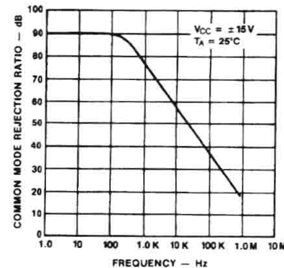
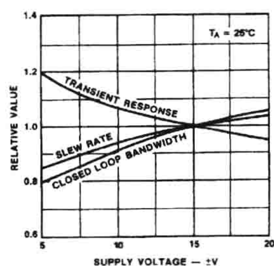


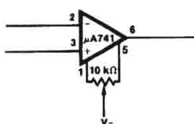
FIGURE 5A.6

Typical Performance Curves (Cont.)

Frequency Characteristics vs Supply Voltage for μA741C/E

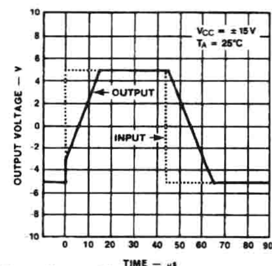


Voltage Offset Null Circuit for μA741C/E

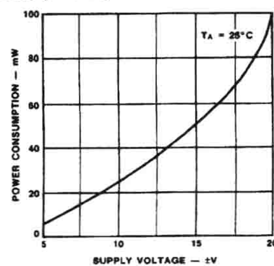


Lead numbers are shown for metal package only

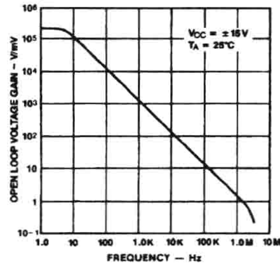
Voltage Follower Large Signal Pulse Response for μA741C/E



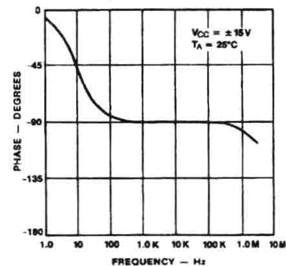
Power Consumption vs Supply Voltage



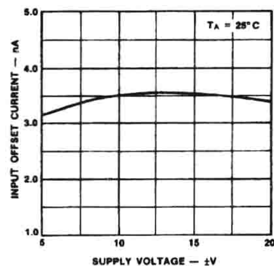
Open Loop Frequency Response



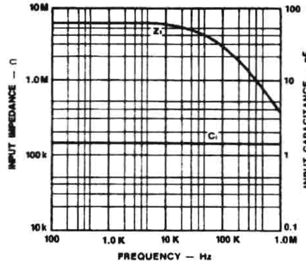
Open Loop Phase Response vs Frequency



Input Offset Current vs Supply Voltage



Input Impedance and Input Capacitance vs Frequency



Output Resistance vs Frequency

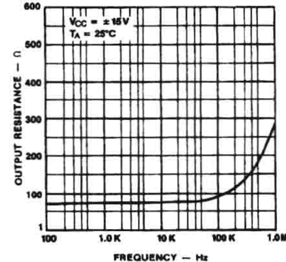
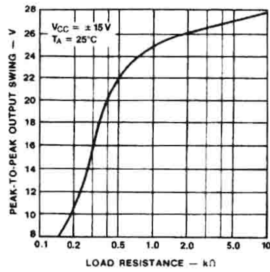


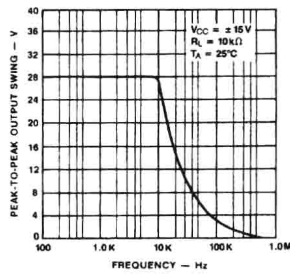
FIGURE 5A.7

Typical Performance Curves (Cont.)

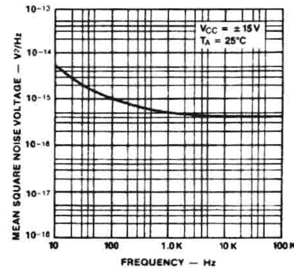
Output Voltage Swing vs Load Resistance



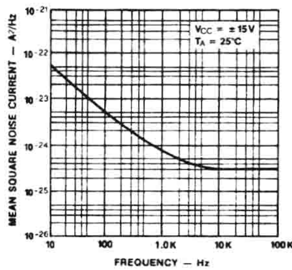
Output Voltage Swing vs Frequency



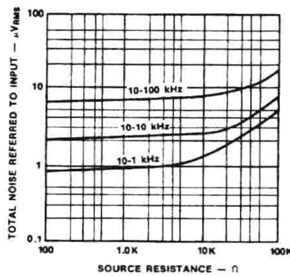
Input Noise Voltage vs Frequency



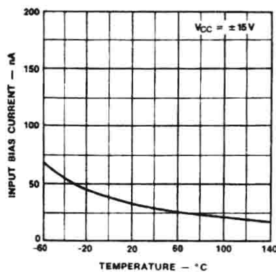
Input Noise Current vs Frequency



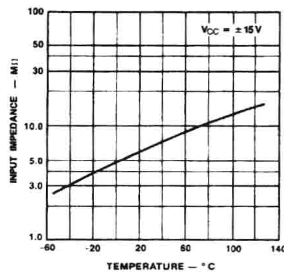
Broadband Noise for Various Bandwidths



Input Bias Current vs Temperature for μA741/A



Input Impedance vs Temperature for μA741/A



Short Circuit Current vs Temperature for μA741/A

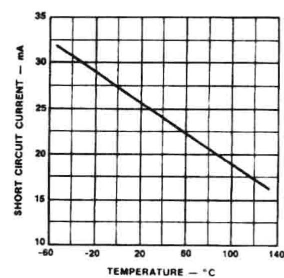
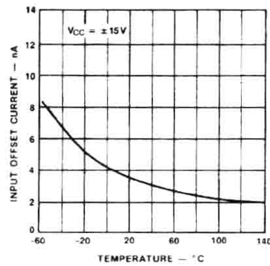


FIGURE 5A.8

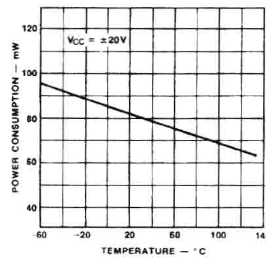
μA741

Typical Performance Curves (Cont.)

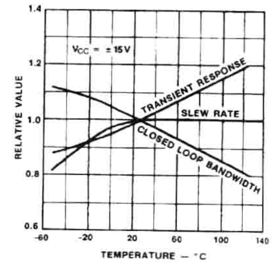
Input Offset Current vs Temperature for μA741/A



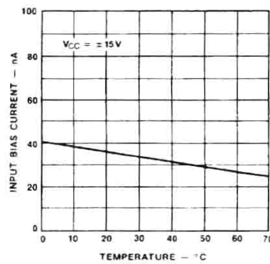
Power Consumption vs Temperature for μA741/A



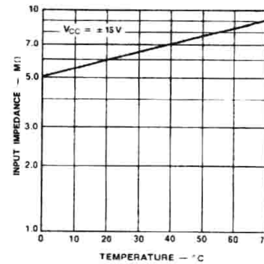
Frequency Characteristics vs Temperature for μA741/A



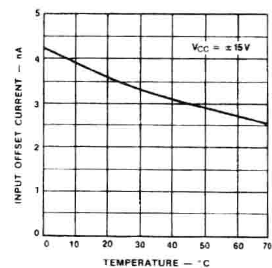
Input Bias Current vs Temperature for μA741C/E



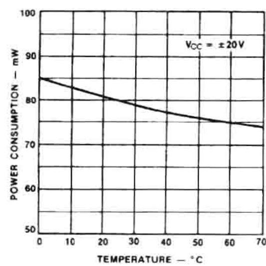
Input Impedance vs Temperature for μA741C/E



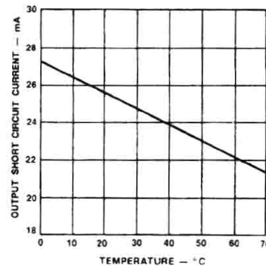
Input Offset Current vs Temperature for μA741C/E



Power Consumption vs Temperature for μA741C/E



Short Circuit Current vs Temperature for μA741C/E



Frequency Characteristics vs Temperature for μA741C/E

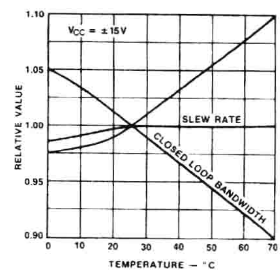


FIGURE 5A.9

6

DYNAMIC OP AMP LIMITATIONS

到目前为止，我们假设运放在任何频率下都具有很高的开环增益。然而，实际的运放只能在直流达到一定频率内提供高增益，超过一定频率后运放的增益就会下降，同时输入与输出之间也会存在延时。这些限制对于电路的闭环特性有很大影响：它们会影响电路的频率响应和瞬态响应速度，改变输入、输出阻抗大小。这一章将要学习单位增益频率 f_t、增益带宽积（GBP）、闭环带宽 f_B、全功率带宽（FPB）、上升时间 t_R、转换速率（SR）和建立时间 t_S，以及这些特性是如何影响一些熟知电路（如四种基本类型放大器、滤波器等）的响应和端口阻抗的。还将讨论一种专门针对高速应用设计的运算放大器——电流反馈放大器（CFA）。

由于在数据手册中频率响应是基于周期频率 f 给出的，所以在这里讨论时将采用这种频率而不是角频率 ω。这两种频率很容易通过计算相互转换 $\omega\leftrightarrow2\pi f$。另外，通过 $jf\leftrightarrow s/2\pi$ 也很容易建立频率响应 $H(jf)$ 和 s 域的对应关系。

运放的开环响应 $a(jf)$ 可能会很复杂，这个将会在第 8 章概括性地讨论。本章将重点介绍一种既特殊又常见的例子，即带有内部补偿的运放。为了使运放稳定，防止产生不需要的振荡，这种运放在内部嵌入了一些片上补偿元件。大多数运放都需要补偿使得开环响应 $a(jf)$ 仅由一个低频极点控制。

本章重点

本章开始时会讨论带有内部补偿运放的开环频率响应以及开环频率响应对环路增益和所有闭环参数的影响。本章将借助简便的图形技巧促进对第 1 章所介绍的 4 种反馈拓扑的形象化理解，包括开环增益、闭环增益和输入/输出阻抗。

接下来会研究电阻运算放大器电路的瞬态响应，包括小信号下运放的线性工作以及大信号下运放非线性工作和转换速率。本章中会大量运用 PSpice 软件来演示运放的频率和瞬态响应。

然后，会对积分器进行详细研究，因为积分器已经成为电路的重要部分。我们需要特别注意积分器的增益和相位误差，尤其是在状态变量和双二阶滤波器的应用中，它们可能会导致不稳定。在此，会引入有源补偿和无源补偿的概念。

积分器的研究是为研究更加复杂的滤波器做准备的，从一阶滤波器着手，到二阶滤波器和滤波器构成模块（如一般的阻抗变换器）。开环增益的衰减对滤波器的性能有很大影响，然而通过预失真补偿使性能达到预想值也是有可能的。再次证明 PSpice 是个强大的处理工具，因为它可以将实际滤波器与理想滤波器的偏离和预失真曲线形象化。

本章最后会介绍电流反馈放大器，这放在最后讲是因为我们需要本章的分析工具来全面评价这种放大器固有的快速动态性。

6.1 OPEN-LOOP FREQUENCY RESPONSE

The most common open-loop response is the *dominant-pole response*, a popular example of which is the 741 response displayed in Fig. 6.1. As we shall see in great detail in Chapter 8, this type of response is designed to prevent oscillation in negative-feedback operation. To understand the basis of the dominant-pole response, refer to Fig. 6.2, which provides a block diagram of the three-stage op amp circuit of Fig. 5.1. Here g_{m1} is the transconductance gain of the first stage, and $-a_2$ is the voltage gain of the second stage, which is an inverting stage. Moreover, R_{eq} and C_{eq} represent the net equivalent resistance and capacitance between the node common to the first and second stage, and ground.

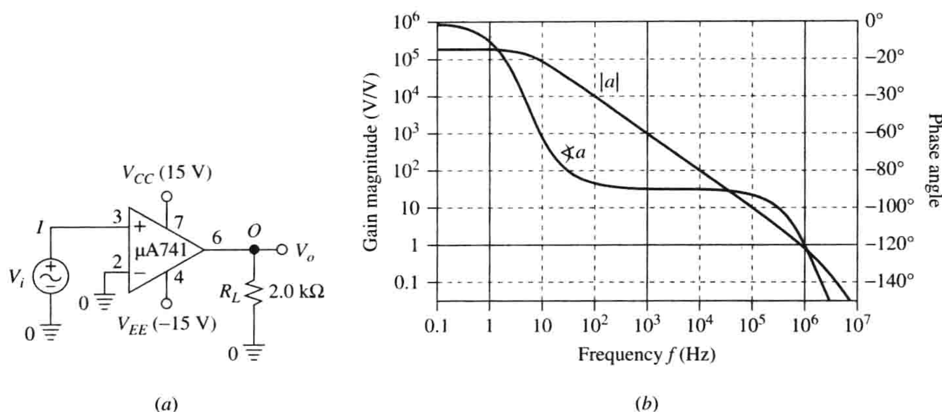


FIGURE 6.1
(*a*) Using the 741 macromodel to plot (*b*) the open-loop frequency response of the 741 op amp. Gain magnitude is plotted as `V(O)/V(I)`, and is in V/V as marked at the left; phase angle is plotted as `P(V(O)/V(I))`, and is in degrees as marked at the right.

At low frequencies, where C_c acts as an open circuit, we have $v_O = 1 \times (-a_2) \times (-R_{\text{eq}} i_{O1}) = g_{m1} R_{\text{eq}} a_2 (v_P - v_N)$. The low-frequency gain, called the *dc gain* and denoted as a_0, is thus

$$a_0 = g_{m1} R_{\text{eq}} a_2 \tag{6.1}$$

As we know, this is a fairly large number. For the 741 op amp we shall assume the following working values: $g_{m1} = 189\ \mu\text{A/V}$, $R_{\text{eq}} = 1.95\ \text{M}\Omega$, and $a_2 = 544$ V/V. Substituting into Eq. (6.1) yields the familiar typical value $a_0 = 200$ V/mV, or 106 dB.

Increasing the operating frequency will bring the impedance of C_{eq} into play, causing gain to roll off with frequency because of the low-pass filter action provided by R_{eq} and C_{eq}. Gain starts to roll off at the frequency f_b that makes $|Z_{C_{\text{eq}}}| = R_{\text{eq}}$, or $1/(2\pi f_b C_{\text{eq}}) = R_{\text{eq}}$. This frequency, called the *dominant-pole frequency,* is thus

$$f_b = \frac{1}{2\pi R_{\text{eq}} C_{\text{eq}}} \tag{6.2}$$

From the data sheets we find that the 741 op amp has typically $f_b = 5$ Hz, indicating a dominant pole at $s = -2\pi f_b = -10\pi$ Np/s. Such a low-frequency pole requires

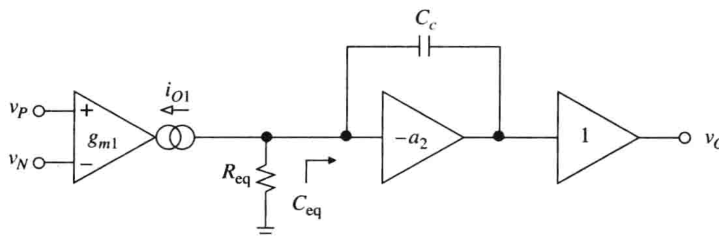


FIGURE 6.2
Simplified op-amp block diagram.

that for a given R_{eq}, C_{eq} be suitably large. For the 741 op amp $C_{\text{eq}} = 1/(2\pi f_b R_{\text{eq}}) = 1/(2\pi 5 \times 1.95 \times 10^6) = 16.3$ nF. The on-chip fabrication of such a large capacitance would be prohibitive in terms of the chip area needed. This drawback is ingeniously avoided by starting out with an acceptable value for C_c, and then exploiting the multiplicative property of the Miller effect to increase its effective value to $C_{\text{eq}} = (1 + a_2)C_c$. The 741 uses $C_c = 30$ pF to achieve $C_{\text{eq}} = (1 + 544)30 = 16.3$ nF.

Closer inspection of Fig. 6.1*b* reveals the presence of additional high-frequency poles because of the steeper rate of magnitude descent and greater phase shift at high frequencies (see Problem 6.1). Should this shift reach $-180°$, feedback would turn from negative to positive, posing the risk of unwanted oscillations. The reason for positioning the dominant pole frequency at such a low value (5 Hz for the 741 op amp) is precisely to ensure that at the frequency of $-180°$ phase shift the gain has dropped *well below unity* so as to render the op amp incapable of sustaining oscillations (more on this in Chapter 8).

Single-Pole Open-Loop Gain

As we proceed, we shall make the simplifying assumption that the open-loop gain $a(s)$ possesses just a *single pole*; this, both to facilitate our mathematical manipulation and to help us develop a basic feel for the effect of the gain rolloff on the closed-loop parameters. Such a gain shall be expressed in the form

$$a(s) = \frac{a_0}{1 + s/\omega_b} \tag{6.3a}$$

where s is the complex frequency, a_0 is the *open-loop dc gain*, and $-\omega_b$ is the s-plane pole location. Alternatively, we express gain in terms of the frequency f as

$$a(jf) = \frac{a_0}{1 + jf/f_b} \tag{6.3b}$$

where j is the imaginary unit ($j^2 = -1$), and $f_b = \omega_b/(2\pi)$ is the *open-loop –3-dB frequency*, also called the *open-loop bandwidth*. We calculate gain magnitude and phase as

$$|a(jf)| = \text{mag}\, a(jf) = \frac{a_0}{\sqrt{1 + (f/f_b)^2}} \tag{6.4a}$$

$$\sphericalangle\, a(jf) = \text{ph}\, a(jf) = -\tan^{-1}(f/f_b) \tag{6.4b}$$

Magnitude is plotted in Fig. 6.3*a*. Shown in Fig. 6.3*b* is an op amp model suitable for basic PSpice simulations. The model uses PSpice's Laplace block to simulate Eq. (6.3*a*) with $a_0 = 10^5$ V/V and $\omega_b = 2\pi(10\text{ Hz})$, as well as $r_d = 1$ MΩ and $r_o = 100$ Ω. Once we have gained sufficient insight using this model, we can always replace it with the macromodel of the op amp we intend to use, and thus investigate any higher-order effects (compared to a macromodel, the simplified model offers the advantage that we can easily alter its parameters as needed).

The gain is high and approximately constant only from dc up to f_b. Past f_b it rolls off at the rate of -20 dB/dec, until it drops to 0 dB (or 1 V/V) at $f = f_t$. This frequency is called the *unity-gain frequency*, or also the *transition frequency*,

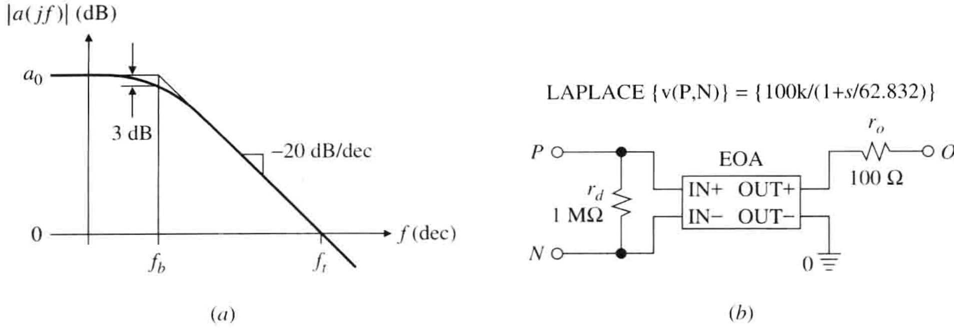


FIGURE 6.3
(*a*) Single-pole open-loop gain, and (*b*) basic PSpice model to simulate an op amp having $a_0 = 10^5$ V/V, $f_b = 10$ Hz, $f_t = 1$ MHz, $r_d = 1$ MΩ, and $r_o = 100$ Ω.

because it marks the transition from amplification (positive decibels) to attenuation (negative decibels). Imposing $1 = a_0/\sqrt{1 + (f_t/f_b)^2}$ in Eq. (6.4a) and using the fact that $f_t \gg f_b$, we get

$$f_t = a_0 f_b \tag{6.5}$$

The 741 op amp has typically $f_t = 200{,}000 \times 5 = 1$ MHz. We wish to emphasize the following special cases:

$$a(jf)\big|_{f \ll f_b} \to a_0 \,\underline{/0^\circ} \tag{6.6a}$$

$$a(jf)\big|_{f = f_b} = \frac{a_0}{\sqrt{2}} \,\underline{/-45^\circ} \tag{6.6b}$$

$$a(jf)\big|_{f \gg f_b} \to \frac{f_t}{f} \,\underline{/-90^\circ} \tag{6.6c}$$

We observe that over the frequency region $f \gg f_b$ the op amp behaves as an integrator, and that its *gain-bandwidth product,* defined as GBP $= |a(jf)| \times f$, is constant

$$\text{GBP} = f_t \tag{6.7}$$

For this reason, op amps with dominant-pole compensation are also referred to as *constant-GBP op amps:* increasing (or decreasing) f by a given amount in the region of integrator behavior will decrease (or increase) $|a|$ by the same amount. This can be exploited to estimate gain at any frequency above f_b. Thus, at $f = 100$ Hz, the 741 has $|a| = f_t/f = 10^6/10^2 = 10{,}000$ V/V; at $f = 1$ kHz, it has $|a| = 1000$ V/V; at $f = 10$ kHz, it has $|a| = 100$ V/V; at $f = 100$ kHz, it has $|a| = 10$ V/V, and so forth (see Fig. 6.1*b*). Browsing through the manufacturers' websites will reveal quite a few op amp families with a gain response of the type of Fig. 6.1*b*. Most general-purpose types tend to have GBPs between 500 kHz and 20 MHz, with 1 MHz being one of the most frequent values. However, for wideband applications, op amp types are available with much higher GBPs. Current-feedback amplifiers, to be discussed in Section 6.7, are an example.

Though a_0 and f_b may be useful for mathematical manipulations, in practice they are ill-defined parameters because so are R_{eq} and a_2, due to manufacturing process

variations. We shall instead focus on the unity-gain frequency f_t, which turns out to be a more predictable parameter. To justify this claim, we note that at high frequencies the circuit of Fig. 6.2 yields $V_o \cong 1 \times Z_{C_c} I_{o1} = [1/(j2\pi f C_c)]g_{m1} \times (V_p - V_n)$, or $a = g_{m1}/(j2\pi f C_c)$. Comparing with Eq. (6.6*c*) gives

$$f_t = \frac{g_{m1}}{2\pi C_c} \tag{6.8a}$$

As discussed in connection with Eq. (5.7), $g_{m1} = I_A/(4V_T)$. Substituting into Eq. (6.8a) gives, for the 741 op amp, the alternative expression

$$f_t = \frac{I_A}{8\pi V_T C_c} \tag{6.8b}$$

It is possible to design for reasonably stable and predictable values of I_A and C_c, thus resulting in a dependable value for f_t. For the 741, $f_t = (19.6 \times 10^{-6})/(8\pi \times 0.026 \times 30 \times 10^{-12}) = 1$ MHz.

Graphical Visualization of the Loop Gain T

In Chapter 1 we learned that even though the op amp is a *voltage* amplifier, through negative feedback it can function also as a *current*, *transresistance*, and *transconductance* amplifier. Yet, the op amp responds to voltage, regardless of the feedback topology in use. In fact, the loop gain T is obtained as the return ratio of voltages, and T is an intrinsic loop parameter that is independent of the type and location of the input and output signals. The frequency plot of $a(jf)$ is usually available from the data sheets, so we seek a way to visualize also the frequency plot of $T(jf)$ in relation to that of $a(jf)$. To this end, we express the loop gain in the form

$$T(jf) = a(jf)\beta(jf) \tag{6.9}$$

where $\beta(jf)$ is found by (*a*) setting all input sources to zero, (*b*) breaking the loop at the output of the op amp's dependent source $a(jf)V_d$, (*c*) applying an ac test voltage V_t downstream of the dependent source, (*d*) finding V_d, and (*e*) finally letting

$$\beta(jf) = -\frac{V_d}{V_t} \tag{6.10}$$

(Alternatively, β can be found merely as $\beta = T/a$.) Just as T should not be confused with L of Section 1.6, β should not be confused with b, though the two parameters may coincide in some cases. In particular, while it is always true that $A_{\text{ideal}} = 1/b$, in general $A_{\text{ideal}} \neq 1/\beta$. (To prevent confusion, avoid using $A_{\text{ideal}} = 1/b$, and always calculate A_{ideal} as the ratio of the output signal to the input signal in the limit $a \to \infty$.) When necessary to distinguish between the two parameters, we shall refer to β as the *return-ratio feedback factor*, and to b as the *two-port feedback factor*.

Rewriting Eq. (6.9) as $T = a/(1/\beta)$ allows us to write $|T|_{\text{dB}} = 20\log_{10}|T| = 20\log_{10}|a| - 20\log_{10}(1/\beta)$, or

$$|T|_{\text{dB}} = |a|_{\text{dB}} - |1/\beta|_{\text{dB}} \tag{6.11a}$$

$$\measuredangle T = \measuredangle a - \measuredangle(1/\beta) \tag{6.11b}$$

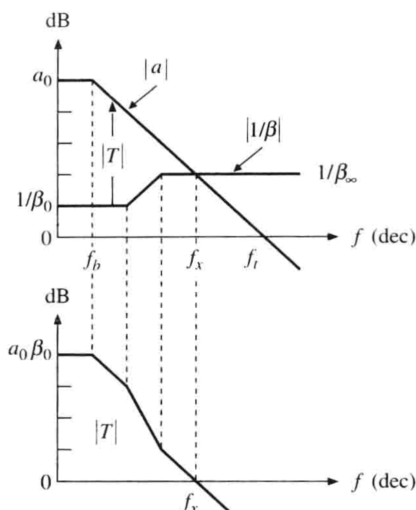


FIGURE 6.4
In a Bode plot, the loop gain $|T|$ is the difference between the $|a|$ and $|1/\beta|$ curves.

indicating that the Bode plots of T can be found graphically as the *difference* between the individual plots of a and $1/\beta$.

Figure 6.4 depicts the magnitude plot. To construct it, we first obtain the open-loop curve from the data sheets. Next, we find β using the techniques of Section 1.7, take its reciprocal $1/\beta$, and then plot $|1/\beta|$. Since usually $|\beta| \leq 1$ V/V, or $|\beta| \leq 0$ dB, it follows that $|1/\beta| \geq 1$ V/V, or $|1/\beta| \geq 0$ dB; that is, the $|1/\beta|$ curve extends above the 0-dB axis. This curve will generally have some breakpoints, though in many cases it is flat. As shown, its low-frequency and high-frequency asymptotes are denoted as $|1/\beta_0|$ and $|1/\beta_\infty|$. Finally, we visualize $|T|$ as the *difference* between the $|a|$ and $|1/\beta|$ curves. The $|T|$ curve is shown explicitly at the bottom, but you should learn to visualize it directly from the diagram at the top.

The frequency f_x at which the two curves meet is called the *crossover frequency*. Clearly, $|T(jf_x)|_{\text{dB}} = 0$ dB, or $|T(jf_x)| = 1$. In the example shown, for $f \ll f_x$ we have $|T| \gg 1$, indicating a closed-loop behavior nearly ideal there. However, for $f > f_x$ we have $|T|_{\text{dB}} < 0$ dB, or $|T| < 1$, indicating a significant departure from the ideal. Thus, the useful frequency range for the op amp circuit is to the left of f_x. In Chapter 8 we shall find that $\measuredangle T(jf_x)$, the phase angle of T at f_x, determines whether a circuit is stable as opposed to oscillatory.

6.2 CLOSED-LOOP FREQUENCY RESPONSE

The fact that the loop gain T is frequency-dependent will make the closed-loop response A depend on frequency even when A_{ideal} is designed to be frequency-independent, as in the case of purely resistive feedback. To stress this fact, we

rewrite Eq. (1.72) as

$$A(jf) = \frac{A_{\text{ideal}}}{1 + 1/T(jf)} + \frac{a_{\text{ft}}}{1 + T(jf)} \tag{6.12}$$

To develop an initial feel, let us assume the feedthrough component to be negligible, and let us express $A(if)$ in the insightful form

$$A(jf) \cong A_{\text{ideal}} D(jf) \tag{6.13a}$$

where

$$D(jf) = \frac{1}{1 + 1/T(jf)} \tag{6.13b}$$

is called the *discrepancy function* because it gives a measure of how close the gain $A(jf)$ is to ideal. The deviation of $D(jf)$ from $1 \angle 0^\circ$ is now specified in terms of two parameters, namely, the *magnitude error*

$$\epsilon_m = \left| \frac{1}{1 + 1/T(jf)} \right| - 1 \tag{6.14a}$$

and the *phase error*

$$\epsilon_\phi = -\sphericalangle[1 + 1/T(jf)] \tag{6.14b}$$

Using Eqs. (6.3) and (6.9), expanding and simplifying, we get

$$D(jf) = \frac{1}{1 + \dfrac{1 + jf/f_b}{a_0\beta}} = \frac{1}{1 + \dfrac{1}{a_0\beta}} \times \frac{1}{1 + \dfrac{jf}{(1 + a_0\beta) f_b}}$$

that is, the discrepancy function is a *low-pass function*

$$D(jf) = \frac{D_0}{1 + jf/f_B} \tag{6.15a}$$

with a *dc value* of

$$D_0 = \frac{1}{1 + 1/(a_0\beta)} \cong 1 \tag{6.15b}$$

and a *–3-db frequency* of

$$f_B = (1 + a_0\beta) f_b \cong a_0\beta f_b = \beta f_t \tag{6.15c}$$

where Eq. (6.5) has been used. Combining Eqs. (6.13) and (6.15), we put the closed-loop gain in the generalized form

$$A(jf) \cong A_0 \frac{1}{1 + jf/f_B} \tag{6.16a}$$

where

$$A_0 = A_{\text{ideal}} D_0 \cong A_{\text{ideal}} \qquad f_B \cong \beta f_t \tag{6.16b}$$

We make the important observation that negative feedback, while reducing gain from a_0 toA_0 by the amount of feedback $1 + a_0\beta$, also *expands* the frequency bandwidth by the same amount, from f_b to f_B. This *gain-bandwidth tradeoff* is yet another important benefit of negative feedback!

Plotting the Closed-Loop Response $|A(jf)|$

By the constancy of the gain-bandwidth product, the crossover frequency in Fig. 6.4 must be such that $(1/\beta) \times f_x = 1 \times f_t$, or $f_x = \beta f_t$, so f_x and f_B coincide, by Eq. (6.16*b*). This offers us a graphical means for constructing the Bode plot of $|A(jf)|$: first, plot the $|1/\beta|$ curve on the plot of $|a|$ supplied by the manufacturer and read out f_B as the crossover frequency between the two curves; then, sketch the low-frequency plot of $|A_{\text{ideal}}|$, and impose a pole breakpoint at $f = f_B$. Let us illustrate with some examples.

The Noninverting and Inverting Amplifiers

The noninverting amplifier of Fig. 6.5*a* admits the loop gain of Eq. (1.76), which we recycle to get the feedback factor as $\beta = T/a$, or

$$\beta = \frac{1}{1 + R_2/R_1 + (R_1 + r_o)/r_d + r_o/R_1} \tag{6.17a}$$

In a well-designed circuit the feedback resistances are much lower than r_d and much higher than r_o, so we approximate the above expression as

$$\beta \cong \frac{1}{1 + R_2/R_1} \tag{6.17b}$$

It is now an easy task to draw the Bode plot of Fig. 6.5*b* with $A_0 \cong 1 + R_2/R_1$ and $f_B \cong f_t/(1 + R_2/R_1)$.

EXAMPLE 6.1. A 741 op amp is configured as a noninverting amplifier with $R_1 = 2\ \text{k}\Omega$ and $R_2 = 18\ \text{k}\Omega$. Find (*a*) the 1% magnitude error and (*b*) the 5° phase error bandwidths, defined, respectively, as the frequency ranges over which $|\epsilon_m| \le 0.01$ and $|\epsilon_\phi| \le 5°$.

Solution.

(*a*) We have $\beta = 0.1$ V/V, so $f_B = \beta f_t = 100$ kHz. By Eq. (6.14*a*), $\epsilon_m = 1/\sqrt{1 + (f/f_B)^2} - 1$. Imposing $|\epsilon_m| \le 0.01$ yields $1/\sqrt{1 + (f/10^5)^2} \ge 0.99$, or $f \le 14.2$ kHz.

(*b*) By Eq. (6.14*b*), $\epsilon_\phi = -\tan^{-1}(f/f_B)$. Imposing $|\epsilon_\phi| \le 5°$ gives $\tan^{-1}(f/10^5) \le 5°$, or $f \le 8.75$ kHz.

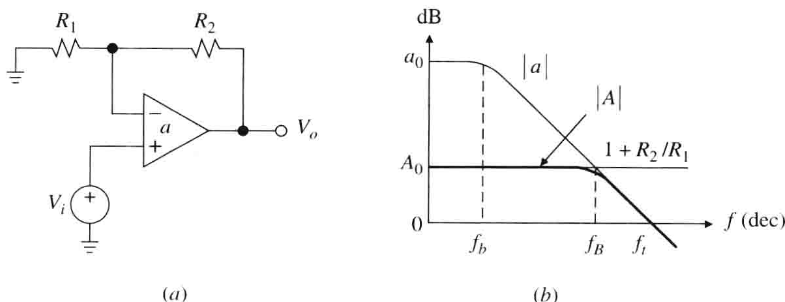


FIGURE 6.5
(*a*) The noninverting amplifier and (*b*) its frequency response.

The gain-bandwidth product for the noninverting amplifier is GBP $= A_0 \times f_B$, or

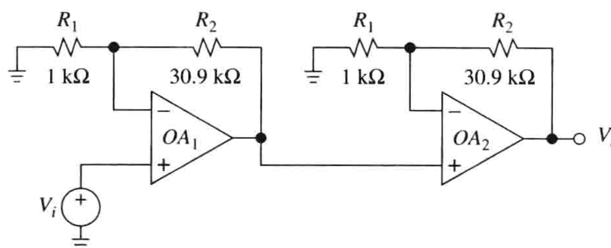
$$\text{GBP}_{\text{noninv}} \cong f_t \qquad (6.18)$$

indicating a *gain-bandwidth tradeoff*. For instance, a 741 op amp configured for $A_0 = 1000$ V/V will have $f_B = f_t/A_0 = 10^6/10^3 = 1$ kHz. Reducing A_0 by a decade, to 100 V/V, will increase f_B also by a decade, to 10 kHz. The amplifier with the lowest gain has also the widest bandwidth: this is the voltage follower, for which $A_0 = 1$ V/V and $f_B = f_t = 1$ MHz. It is apparent that f_t represents a figure of merit for op amps. The gain-bandwidth tradeoff can be exploited to meet specific bandwidth requirements, as illustrated in the following example.

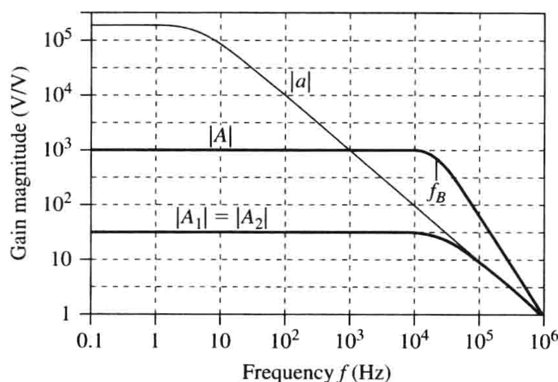
EXAMPLE 6.2. (*a*) Using 741 op amps, design an audio amplifier with a gain of 60 dB. (*b*) Sketch its magnitude plot. (*c*) Find its actual bandwidth.

Solution.

(*a*) Since $10^{60/20} = 10^3$, the design calls for an amplifier with $A_0 = 10^3$ V/V and $f_B \geq$ 20 kHz. A single 741 op amp will not do, because it would have $f_B = 10^6/10^3 =$ 1 kHz. Let us try cascading two noninverting stages with lesser individual gains but wider bandwidths, as depicted in Fig. 6.6*a*. Denoting the individual gains as A_1 and A_2, the overall gain is then $A = A_1 \times A_2$. One can easily prove that the widest bandwidth for A is achieved when A_1 and A_2 are made equal, or $A_{10} = A_{20} = \sqrt{1000} = 31.62$ V/V, or 30 dB. Then, $f_{B1} = f_{B2} = 10^6/31.62 = 31.62$ kHz.



(*a*)



(*b*)

FIGURE 6.6
Cascading two amplifiers and the resulting frequency response $|A|$.

(*b*) To construct the magnitude plot, refer to Fig. 6.1*b* and note that since $A = A_1^2$, we have $|A|_{dB} = 2|A_1|_{dB}$, indicating that the magnitude plot of A is obtained by multiplying that of A_1 by 2, point by point. The plot of $|A_1|$ is in turn obtained via the graphical technique of Fig. 6.5*b*. The final result is shown in Fig. 6.6*b*.

(*c*) Note that at 31.62 kHz both $|A_1|$ and $|A_2|$ are 3 dB below their dc values, making $|A|$ in turn 6 dB below its dc value. The −3-dB frequency f_B is such that $|A(jf_B)| = 10^3/\sqrt{2}$. But, $|A(jf)| = |A_1(jf)|^2 = 31.62^2/[1 + (f/f_B)^2]$. We thus impose

$$\frac{10^3}{\sqrt{2}} = \frac{31.62^2}{1 + [f_B/(31.62 \times 10^3)]^2}$$

to obtain $f_B = 31.62\sqrt{\sqrt{2} - 1} = 20.35$ kHz, which indeed meets the audio bandwidth requirement.

The inverting amplifier of Fig. 6.7*a* admits the same β as its noninverting counterpart, so Eqs. (6.17) still hold. As depicted in Fig. 6.7*b*, f_B still coincides with the crossover frequency of the $|a|$ and $|1/\beta|$ curves, so $f_B \cong f_t/(1+R_2/R_1)$. However, we now have $A_0 \cong -R_2/R_1$, whose magnitude is *less* than that of $1 + R_2/R_1$, so the $|A|$ curve will be shifted downward somewhat.

It is readily seen that the inverting configuration admits

$$\text{GBP}_{\text{inv}} \cong (1 - \beta) f_t \tag{6.19}$$

The biggest difference between the two amplifier types occurs when they are both configured for unity gain: in the noninverting case we use $R_1 = \infty$ and $R_2 = 0$, so $\beta = 1$ and $f_B \cong f_t$, whereas in the inverting case we use $R_1 = R_2$, so $\beta = 1/2$ and $f_B \cong 0.5 f_t$. For high closed-loop gains, β is very small, so the difference between the two GBPs becomes negligible.

The *I*-*V* and *V*-*I* Converters

Using Eq. (6.10) we readily find the feedback factor of the I-V converter of Fig. 6.8*a* to be

$$\beta = \frac{r_d}{r_d + R + r_o} \tag{6.20}$$

which we use in connection with the plot of $|a|$ to obtain $f_B (= \beta f_t)$ as in Fig. 6.8*b* (top). As we know, $A_0 \cong -R$ V/A, so we plot the closed-loop response $|A|$ as shown at the bottom (note that the units of A are different from those of a, thus mandating separate plots).

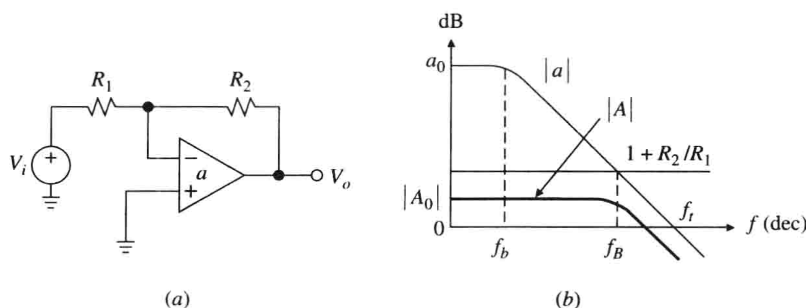


FIGURE 6.7
(*a*) The inverting amplifier and (*b*) its frequency response $|A|$.

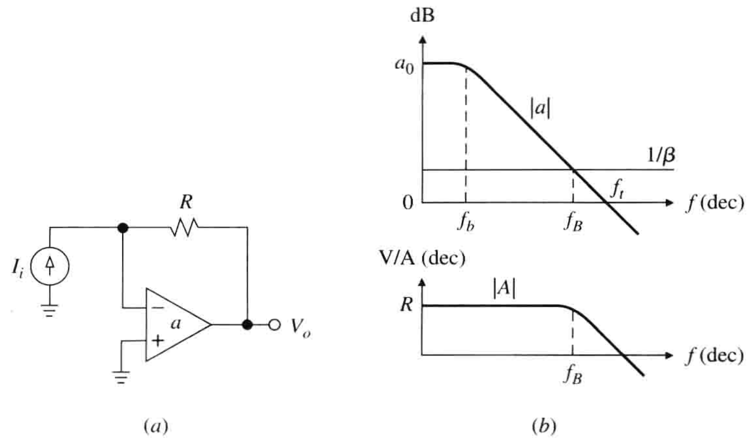


FIGURE 6.8
(*a*) The *I*-*V* converter and (*b*) its frequency response.

The feedback factor of the *V*-*I* converter of Fig. 6.9*a* is

$$\beta = \frac{r_d \| R}{r_d \| R + R_L + r_o} \tag{6.21}$$

Moreover, $A_0 \cong 1/R$ A/V, so the required plots are as in Fig. 6.9*b*.

Before concluding this section, let us take a closer look at the approximations made above. Let us start with the approximation leading to Eq. (6.17*b*), which we examine in the following example.

EXAMPLE 6.3. (*a*) Suppose the op amp of Fig. 6.3*b* is configured as a noninverting amplifier with $R_1 = R_2 = 10\ \text{k}\Omega$. Find A_0 and f_B. (*b*) Repeat, but with $R_1 = R_2 = 1.0\ \text{M}\Omega$. Compare with (*a*) and comment.

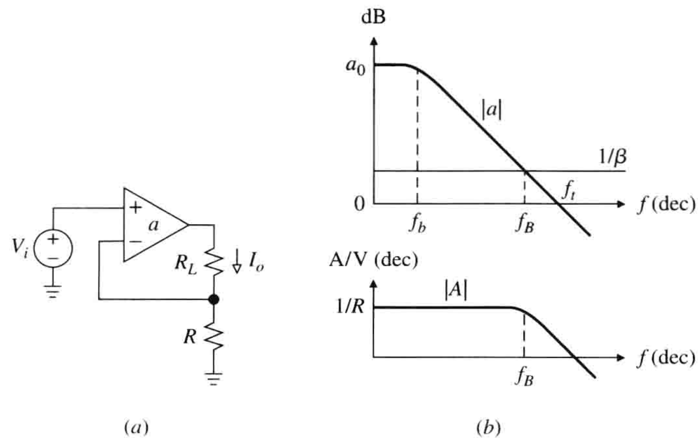


FIGURE 6.9
(*a*) The *V*-*I* converter and (*b*) its frequency response.

Solution.

(*a*) We have $A_0 \cong A_{\text{ideal}} = 1 + R_2/R_1 = 2.0$ V/V. Plugging the given data into Eq. (6.17*a*) gives $\beta = 0.495$, which is only slightly lower than the value of 0.5 given by Eq. (6.17*b*). The op amp has $f_t = a_0 f_b = 10^5 \times 10 = 1$ MHz, so $f_B = \beta f_t = 495$ kHz ($\cong 500$ kHz).

(*b*) We still have $A_{\text{ideal}} = 2.0$ V/V. However, Eq. (6.17*a*) now gives $\beta = 0.333$, so $f_B = 333$ kHz. Because of the much larger resistances used, we can no longer ignore loading by r_d. While the low-frequency asymptote remains virtually unchanged at $A_0 \cong 2.0$ V/V, the $1/\beta$ curve is *shifted upward*, in turn decreasing the crossover frequency and, hence, the value of f_B. Note that the two-port feedback factor is $b = 0.5 (= 1/A_{\text{ideal}})$ in both cases, whereas the return-ratio feedback factor β essentially coincides with b in (*a*), but decreases to 0.333 in (*b*). This example should help the reader appreciate similarities and differences between b and β.

Finally, let us examine the impact of ignoring the feedthrough component in Eq. (6.12). This is certainly acceptable in the case of the noninverting amplifier because feedthrough occurs via r_d, which is large. Not necessarily so in the inverting case, where feedthrough occurs via the usually much smaller resistances of the feedback network. The following example will give an idea.

EXAMPLE 6.4. Suppose the op amp of Fig. 6.3*b* is configured as an inverting amplifier with $R_1 = R_2 = 500\ \Omega$. Use PSpice to display the magnitude Bode plots of a, $1/\beta$, and A. Discuss all salient features (asymptotic values, breakpoints), and confirm via hand calculations.

Solution. The circuit and the plots are shown in Fig. 6.10. Clearly, $A_{\text{ideal}} = -R_2/R_1 = -1.0$ V/V, yielding the low-frequency asymptote $|A_0| \cong 1$ V/V $= 0$ dB. The values of R_1 and R_2 have been chosen deliberately low both to load down r_o and to cause noticeable feedthrough. Equation (6.17*a*) gives $\beta = 0.4544$, so $f_B = \beta f_t \cong 454$ kHz, which is

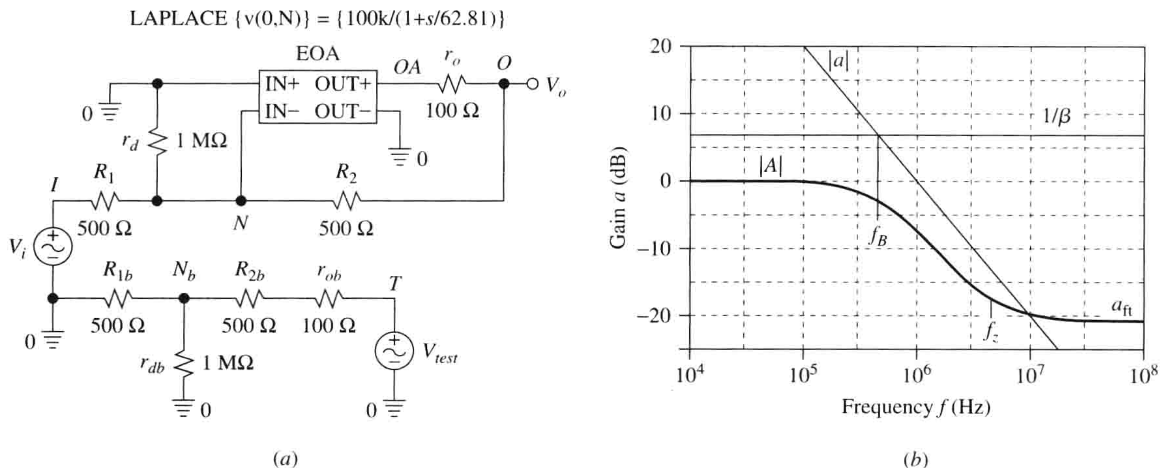


FIGURE 6.10
(*a*) The inverting amplifier of Example 6.4 and (*b*) its frequency response. The feedback network for plotting $1/\beta$ is repeated at the bottom of (*a*), with its elements identified by subscript *b*. The $|A|$ trace is plotted as `DB(V(O)/V(I))`, the $|a|$ trace as `DB(V(OA)/(-V(N)))`, and the $1/\beta$ trace as `DB(V(T)/V(Nb))`.

less than 500 kHz as given by Eq. (6.17*b*). Because of feedthrough, $|A|$ exhibits a high-frequency asymptote of a_{ft}. By Eq. (1.75*b*) we can write $a_{ft} = \beta(r_o/R_1) = 0.0909$ V/V $= -20.8$ dB. Beside the pole frequency f_B, $A(jf)$ exhibits also a zero frequency f_z. Exploiting the constancy of the GBP, we write $A_0 \times f_B = a_{ft} \times f_z$, which gives $f_z = A_0 f_B / a_{ft} \cong 5$ MHz. All calculated data match the PSpice data, which were measured directly on the traces using PSpice's cursor.

6.3 INPUT AND OUTPUT IMPEDANCES

As we know, the closed-loop input/output characteristics of an op amp circuit are governed by the loop gain via Blackman's formula. Since the loop gain is frequency-dependent, so are the terminal characteristics, which we shall henceforth refer to as *impedances*. We thus rephrase Eq. (1.79) as

$$Z = z_0 \frac{1 + T_{sc}}{1 + T_{oc}} \qquad z_0 = \lim_{a \to 0} Z \tag{6.22}$$

The examples of Section 1.7 have revealed that series-type ports tend to have $T_{oc} = 0$ and $T_{sc} = T$, so

$$Z_{se} = z_0(1 + T) \tag{6.23a}$$

whereas shunt-type ports tend to have $T_{sc} = 0$ and $T_{oc} = T$, so

$$Z_{sh} = \frac{z_0}{1 + T} \tag{6.23b}$$

If *both* z_0 and β are *frequency-independent*, we can expedite the construction of the Bode plots of $|Z_{se}(jf)|$ and $|Z_{sh}(jf)|$ via the graphical techniques depicted in Fig. 6.11*a*. To this end, we first draw the $1/\beta$ curve on the plot of $|a(jf)|$ and read out the value of f_B. Next, on a separate graph with the vertical axis scaled logarithmically in ohms, we draw the low-frequency asymptotes Z_{se0} and Z_{sh0}, examples of which were calculated in Section 1.6. These asymptotes hold only up to f_b. For $f > f_b$, the loop gain $|T|$ rolls off with frequency, causing in turn $|Z_{se}(jf)|$ to *fall* and $|Z_{sh}(jf)|$ to *rise*. The fall/rise continues up to f_B. For $f > f_B$, the loop gain $|T|$ becomes negligible compared to unity, in turn causing the impedances to settle at their high-frequency asymptotic values $Z_{se\infty}$ and $Z_{sh\infty}$. These values are given by Eq. (1.79*b*). Evidently, the dramatic resistance transformations investigated in Chapter 1 hold only at low frequencies, where the amount of feedback is high. As the gain $a(jf)$ rolls off, the benefits of negative feedback gradually decrease, until they evaporate altogether above f_B.

For calculation purposes, we can express the impedances mathematically as

$$Z_{se}(jf) = Z_{se0} \frac{1 + jf/f_B}{1 + jf/f_b} \qquad Z_{sh}(jf) = Z_{sh0} \frac{1 + jf/f_b}{1 + jf/f_B} \tag{6.24}$$

We observe that *series* impedances tend to behave *capacitively*, and *shunt* impedances *inductively*. In fact, it is good practice to visualize them in terms of their equivalents of Fig. 6.11*b*. At low frequencies C_{eq} acts as an open circuit, so $Z_{se} \to Z_{se0}$, whereas at high frequencies C_{eq} acts as a short circuit, so $Z_{se} \to Z_{se0} \| Z_{se\infty} \cong Z_{se\infty}$ because $Z_{se\infty} \ll Z_{se0}$. By dual reasoning, at high

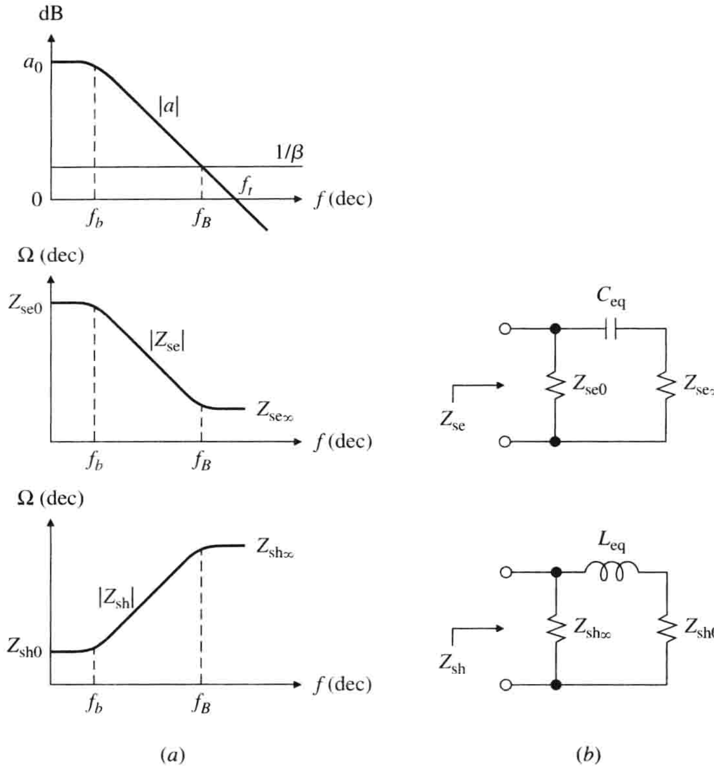


FIGURE 6.11
(*a*) Graphical technique for drawing the Bode plots of series/shunt impedances. (*b*) Approximate equivalent circuits of a series impedance (top) and a shunt impedance (bottom).

frequencies L_{eq} acts as an open circuit, so $Z_{sh} \to Z_{se\infty}$, whereas at low frequencies L_{eq} acts as a short circuit, so $Z_{sh} \to Z_{sh\infty} \| Z_{sh0} \cong Z_{sh0}$ because $Z_{sh0} \ll Z_{sh\infty}$.

Let us use physical insight to derive expressions for C_{eq} and L_{eq}. The effect of C_{eq} starts to be felt as f is raised to f_b, where its impedance becomes equal in magnitude to Z_{se0}, so $1/(2\pi f_b C_{eq}) = Z_{se0}$. By dual reasoning, L_{eq} starts to be felt as f is lowered to f_B, where its impedance becomes equal in magnitude to $Z_{sh\infty}$, so $2\pi f_B L_{eq} = Z_{sh\infty}$. Solving for C_{eq} and L_{eq} we get

$$C_{eq} = \frac{1}{2\pi f_b Z_{se0}} \qquad L_{eq} = \frac{Z_{sh\infty}}{2\pi f_B} \tag{6.25}$$

In Chapter 8 we shall see that the inductive behavior of shunt-type ports may cause instability when the port is terminated on a capacitance, whether intentional or parasitic: such a capacitance tends to form a resonant circuit with L_{eq}, possibly causing intolerable peaking and ringing, or even unwanted oscillation, unless the circuit is properly damped. A familiar example is the stray capacitance of the inverting-input terminal, which tends to destabilize I-V and I-I converters and inverting voltage amplifiers. Another example is the load capacitance arising when a shunt-output amplifier drives a long cable. (Circuit stabilization techniques, called frequency compensation, are addressed in Chapter 8.)

EXAMPLE 6.5. Suppose the op amp of Fig. 6.3*b* is configured as a noninverting amplifier with $R_1 = 2.0\ \text{k}\Omega$ and $R_2 = 18\ \text{k}\Omega$. (*a*) Find the asymptotic values and the break frequencies of the magnitude Bode plot of its input impedance $Z_i(jf)$. What are the element values of its equivalent circuit? (*b*) Repeat, but for the output impedance $Z_o(jf)$.

Solution.

(*a*) We have $\beta \cong 1/10$, so $f_B = \beta f_t \cong 100$ kHz. The input port is of the series type, so $Z_{i0} \cong r_d(1+a_0\beta) = 10^6(1+10^5/10) = 10\ \text{G}\Omega$. By inspection, $Z_{i\infty} \cong r_d = 1\ \text{M}\Omega$. Finally, $C_{\text{eq}} = 1/(2\pi \times 10 \times 10^{10}) = 1.59$ pF. The equivalent circuit consists of 10 GΩ in parallel with the series combination of 1 MΩ and 1.59 pF.

(*b*) The output port is of the shunt type, so $Z_{o0} \cong r_o/(1+a_0\beta) = 10\ \text{m}\Omega$. By inspection, $Z_{o\infty} \cong r_o = 100\Omega$. Finally, $L_{\text{eq}} = Z_{o\infty}/(2\pi \times 10^5) = 159\ \mu\text{H}$. The equivalent circuit consists of 100 Ω in parallel with the series combination of 10 mΩ and 159 μH.

EXAMPLE 6.6. The op amp of Fig. 6.3*b* is configured for current amplification as in Fig. 1.30*b* with $R_1 = 1.0\ \text{k}\Omega$ and $R_2 = 99\ \text{k}\Omega$. Use PSpice to display the magnitude Bode plots of a, $1/\beta$, the closed-loop gain A_i, and the input and output impedances Z_i and Z_o for the case of a short-circuit load. Discuss all salient features (asymptotic values, breakpoints), and verify via hand calculations.

Solution. With reference to Fig. 6.12*a*, we first use the circuit at the top to plot a, $1/\beta$, $A_i = I_o/I_i$, and $Z_i = V_n/I_i$. Then, we suppress the input source and apply a test

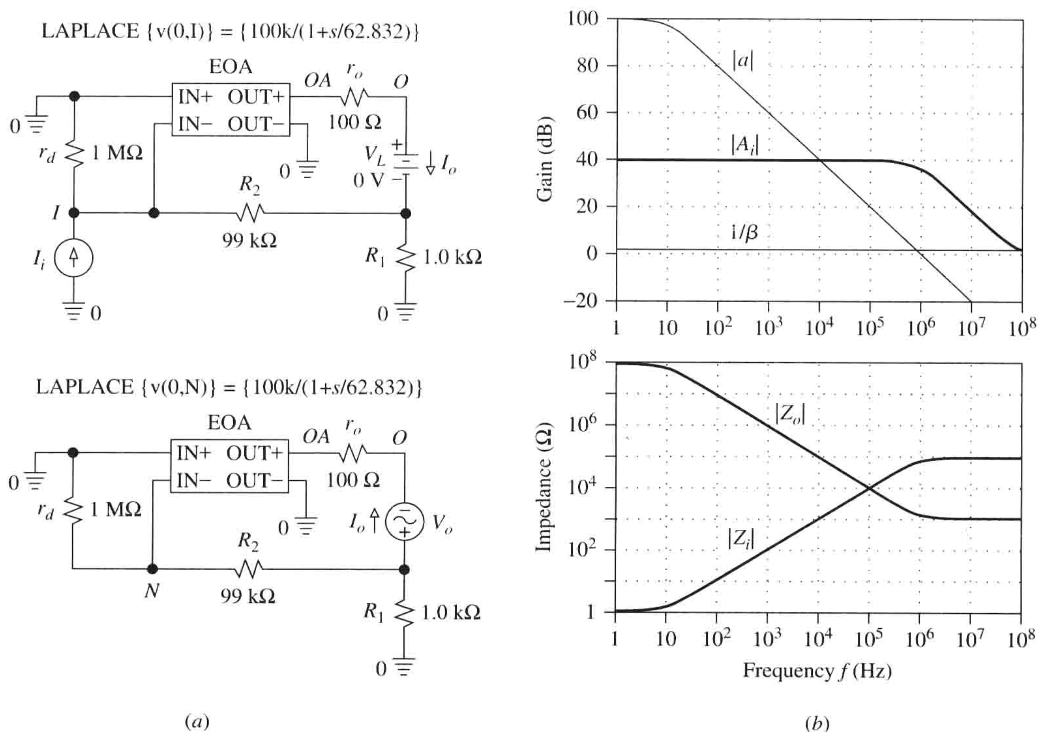


FIGURE 6.12
(*a*) PSpice circuits for the current amplifier of Example 6.6 and (*b*) its Bode plots. The $|Z_i|$ trace is plotted on a log scale as `V(I)/I(Ii)`, and the $|Z_o|$ trace as `V1(Vo)/I(Vo)`.

voltage as in the bottom circuit, and let $Z_o = V_o/I_o$. Equation (1.63) predicts $A_{i(\text{ideal})} = -100$ A/A. Moreover, Eq. (6.10) gives

$$\beta = \frac{r_d}{r_d + R_2} \times \frac{(r_d + R_2) \| R_1}{(r_d + R_2) \| R_1 + r_o} = 0.827$$

so $1/\beta = 1.65$ dB and $f_B = \beta f_t = 827$ kHz. By inspection, $Z_{i\infty} = r_d \| (R_2 + R_1 \| r_o) \cong 90$ kΩ, and $Z_{o\infty} = r_o + R_1 \| (R_2 + r_d) \cong 1.1$ kΩ. Finally, $Z_{i0} = Z_{i\infty}(f_b/f_B) \cong 1.0\Omega$, and $Z_{o0} = Z_{o\infty}(f_B/f_b) \cong 91$ MΩ. All calculated data match PSpice's data. As a final note we observe that because of feedthrough, A_i exhibits a high-frequency asymptote just a bit below 0 dB.

Practical Considerations

The above analysis assumed the open-loop input and output impedances to be purely resistive, $z_d = r_d$ and $z_o = r_o$. However, closer inspection of Fig. 5A.7 indicates that at high frequencies z_d tends to behave capacitively and z_o tends to behave inductively. This behavior is typical of most op amps, and is due primarily to the stray capacitances of the input transistors and to the frequency limitations of the output transistors. Moreover, if the inputs of a practical op amp are tied together and the impedance to ground is measured, the result is the *common-mode input impedance* z_c. In the op amp model of Fig. 6.13, z_c has been split equally between the two inputs in order to yield $(2z_c) \| (2z_c) = z_c$ when they are tied together.

Data sheets usually specify only the resistive portion of these impedances, namely, r_d, r_c, and r_o. For BJT-input op amps, r_d and r_c are typically in the megaohm and gigaohm range, respectively. Since $r_c \gg r_d$, the specification of r_c is often omitted, and only r_d is given. For FET-input devices, r_d and r_c are of the same order of magnitude and in the range of 100 GΩ or higher.

A few manufacturers specify the reactive portions of z_d and z_c, namely, the *differential input capacitance* C_d, and the *common-mode input capacitance* C_c. For example, the AD705 op amp has typically $z_d = r_d \| C_d = (40\ \text{M}\Omega) \| (2\ \text{pF})$ and $z_c = r_c \| C_c = (300\ \text{G}\Omega) \| (2\ \text{pF})$. In general, it is safe to assume values on the order of few picofarads for both C_d and C_c. Though irrelevant at low

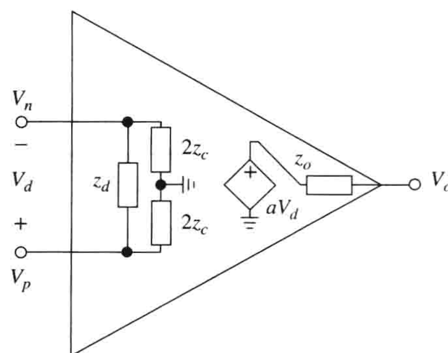


FIGURE 6.13
Modeling the input and output impedances of a practical op amp.

frequencies, these capacitances may cause significant degradation at high frequencies. For instance, at dc the AD705 op amp has $z_c = r_c = 300\ \text{G}\Omega$; however, at 1 kHz, where $Z_{C_c} = 1/(j2\pi \times 10^3 \times 2 \times 10^{-12}) \cong -j80\ \text{M}\Omega$, it has $z_c = (300\ \text{G}\Omega) \parallel (-j80\ \text{M}\Omega) \cong -j80\ \text{M}\Omega$, a drastically reduced magnitude.

It is apparent that the assumptions $z_d = r_d$ and $z_o = r_o$, while a good starting point as we develop a feel for the effects of gain rolloff, may fail to account for certain high-frequency effects. Taking also the reactive components of z_d and z_o into consideration may prove too arduous for hand analysis, so computer simulation using adequate macromodels becomes a necessity in this case.

6.4 TRANSIENT RESPONSE

So far we have investigated the effect of the open-loop dominant pole in the frequency domain. We now turn to the time domain by examining the transient response, that is, the response to an input step as a function of time. This response, like its frequency-domain counterpart, varies with the amount of feedback applied. In the data sheets it is usually specified for unity feedback, that is, for the voltage follower configuration; however, the results can readily be generalized to other feedback factors.

The Rise Time t_R

As we know, the small-signal bandwidth of the voltage follower is f_t, so its frequency response can be written as

$$A(jf) = \frac{1}{1 + jf/f_t} \tag{6.26}$$

indicating a pole at $s = -2\pi f_t$. Subjecting the voltage follower of Fig. 6.14*a* to an input voltage step of sufficiently small amplitude V_m will result in the well-known exponential response

$$v_O(t) = V_m(1 - e^{-t/\tau}) \tag{6.27}$$

$$\tau = \frac{1}{2\pi f_t} \tag{6.28}$$

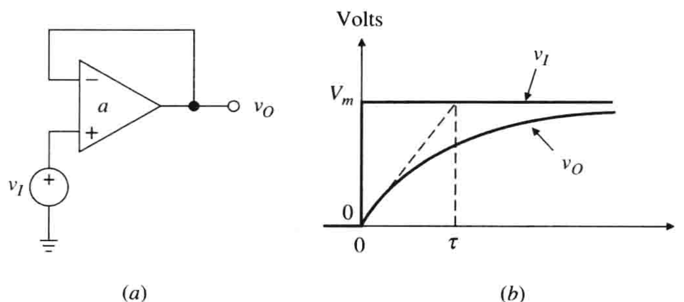


FIGURE 6.14
Voltage follower and its small-signal step response.

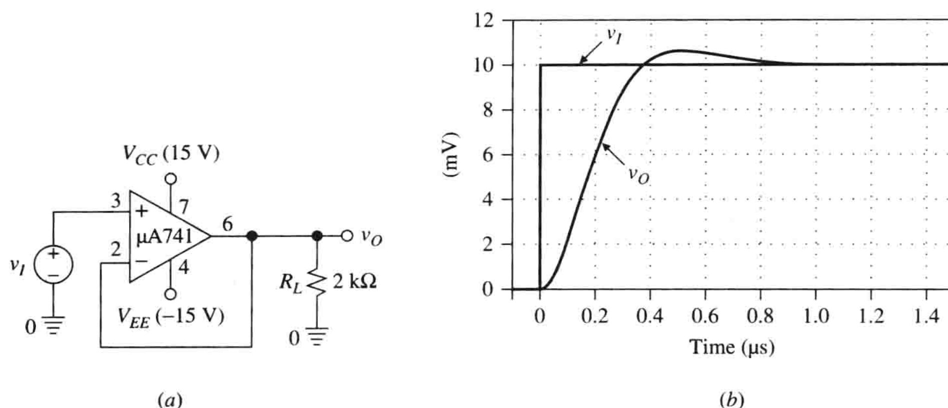


FIGURE 6.15
(*a*) PSpice circuit to display 741 transient responses. (*b*) Small-signal step response.

The time t_R it takes for v_O to swing from 10% to 90% of V_m is called the *rise time,* and it provides an indication of how rapid the exponential swing is. We easily find $t_R = \tau(\ln 0.9 - \ln 0.1)$, or

$$t_R = \frac{0.35}{f_t} \tag{6.29}$$

This provides a link between the frequency-domain parameter f_t and the time-domain parameter t_R; clearly, the higher f_t, the lower t_R.

The 741 op amp has $\tau = 1/(2\pi 10^6) \cong 159$ ns and $t_R \cong 350$ ns. Its transient response, shown in Fig. 5A.6 of Appendix 5.A, is repeated for convenience in Fig. 6.15 using PSpice's 741 macromodel. The small amount of ringing is due to the aforementioned presence of higher-order poles, which we have ignored in our single-pole approximation.

Slew-Rate Limiting

The rate at which v_O changes with time is highest at the beginning of the exponential transition. Using Eq. (6.27), we find $dv_O/dt|_{t=0} = V_m/\tau$, which is also illustrated in Fig. 6.14*b*. If we increase V_m, the rate at which the output slews will have to increase accordingly in order to complete the 10%-to-90% transition within the time t_R. In practice it is observed that above a certain step amplitude the output slope saturates at a constant value called the *slew rate* (SR). The output waveform, rather than an exponential curve, is now a ramp. (As we shall see in greater detail shortly, slew-rate limiting is a nonlinear effect that stems from the limited ability by the internal circuitry to charge or discharge the frequency-compensation capacitance C_c.)

The SR is expressed in volts per microsecond. The data sheets give SR = 0.5 V/μs for the 741C op amp version and SR = 0.7 V/μs for the 741E version. This means that to complete a 1-V output swing, a 741C voltage follower takes approximately (1 V)/(0.5 V/μs) = 2μs. This is confirmed by Fig. 6.16*a*, generated via the PSpice circuit of Fig. 6.15*a*.

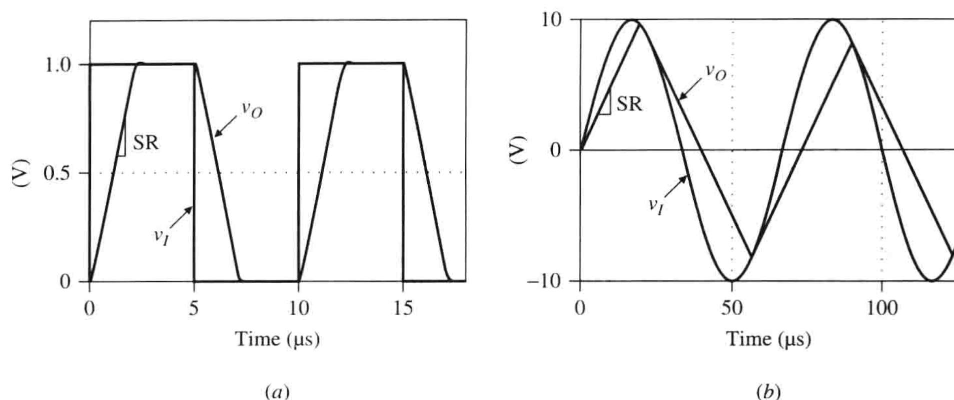


FIGURE 6.16
Slew-rate limited responses of the 741 follower of Fig. 6.15*a* to (*a*) a pulse and (*b*) a sinusoid.

We stress that SR is a nonlinear large-signal parameter, while t_R is a linear small-signal parameter. The critical output-step magnitude corresponding to the onset of slew-rate limiting is such that $V_{om(\text{crit})}/\tau = \text{SR}$. Using Eq. (6.28), this gives

$$V_{om(\text{crit})} = \frac{\text{SR}}{2\pi f_t} \tag{6.30}$$

For the 741C, $V_{om(\text{crit})} = 0.5 \times 10^6/(2\pi \times 10^6) = 80$ mV. This means that as long as the input step is less than 80 mV, a 741C voltage follower responds with an approximately exponential transition governed by $\tau = 159$ ns. However, for a greater input step, the output slews at a constant rate of 0.5 V/μs until it comes within 80 mV of the final value, after which it performs the remainder of the transition in approximately exponential fashion. The above results can be generalized to circuits with $\beta < 1$ by replacing f_t with βf_t.

EXAMPLE 6.7. Suppose a 741 op amp is configured as an inverting amplifier with $R_1 = 10$ kΩ and $R_2 = 20$ kΩ. Use PSpice to display $v_O(t)$ and $v_N(t)$ if $v_I(t)$ is a 0-to-1-V step. Discuss the salient features of each waveform, and verify via hand calculations.

Solution. With reference to Fig. 6.17, we observe that with a gain of −2 V/V, the circuit responds to the 0-to-1 V input step with an output transition from 0 to −2 V. Initially, this transition is slew-rate limited, and only at the end does it become exponential. The slew-rate limited portion takes close to $2/0.5 = 4$ μs to complete. During slew-rate limiting the op amp is unable to influence v_N with all its might, so we use the superposition principle to find

$$v_N = \frac{2}{3}v_I + \frac{1}{3}v_O = \frac{1}{3}v_O + 0.667 \text{ V}$$

The greatest departure of v_N from virtual-ground behavior occurs right at the beginning of the output transition, when v_O is still 0 and v_N thus jumps from 0 to 0.667 V. Only after the op amp stops slew-rate limiting will v_N approach virtual-ground behavior. The final, exponential portion is governed by the time constant $\tau = 1/(2\pi\beta f_t)$. With $\beta \cong 1/3$, we get $\tau = 477$ ns. Moreover, since SR/$(2\pi\beta f_t) = 239$ mV, it follows that slew-rate limiting occurs until v_O comes within 0.239 V of its final value of −2.0 V, the remainder of the transition becoming approximately exponential.

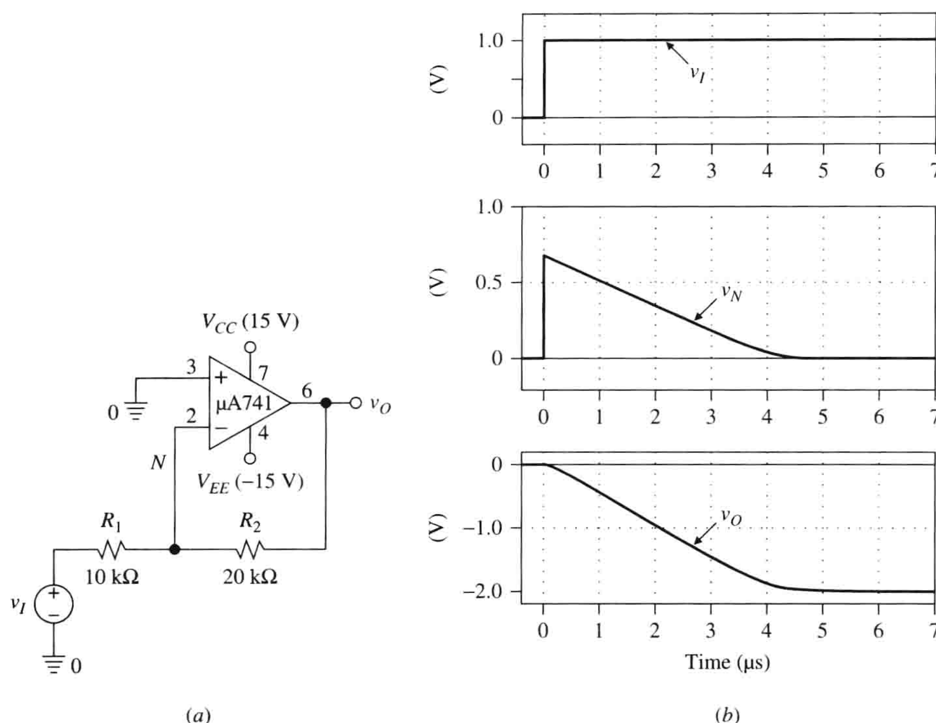


FIGURE 6.17
(a) PSpice circuit of Example 6.7, and (b) its step responses.

Full-Power Bandwidth

The effect of slew-rate limiting is to distort the output signal whenever an attempt is made to exceed the SR capabilities of the op amp. This is illustrated in Fig. 6.16b for a sinusoidal signal. In the absence of slew-rate limiting, the output would be $v_O = V_{om}\sin(2\pi ft)$. Its rate of change is $dv_O/dt = 2\pi f V_{om}\cos(2\pi ft)$, whose maximum is $2\pi f V_{om}$. To prevent distortion, we must require $(dv_O/dt)_{\max} \leq \text{SR}$, or

$$f V_{om} \leq \text{SR}/(2\pi) \tag{6.31}$$

indicating a tradeoff between frequency and amplitude. If we want to operate at high frequencies, then we must keep V_{om} suitably small to avoid slew-rate distortion. In particular, if we want to exploit the full small-signal bandwidth f_t of a 741C voltage follower, then we must keep $V_{om} \leq \text{SR}/(2\pi f_t) \cong 80$ mV. Conversely, if we want to ensure an undistorted output with $V_{om} > V_{om(\text{crit})}$, then we must keep $f \leq \text{SR}/(2\pi V_{om})$. For instance, for an undistorted ac output with $V_{om} = 1$ V, a 741C follower must be operated below $0.5 \times 10^6/(2\pi 1) = 80$ kHz, which is way below $f_t = 1$ MHz.

The *full-power bandwidth* (FPB) is the maximum frequency at which the op amp will yield an undistorted ac output with the largest possible amplitude. This amplitude depends on the particular op amp as well as its power supplies. Assuming

symmetric output saturation values of $\pm V_{sat}$, we can write

$$\text{FPB} = \frac{\text{SR}}{2\pi V_{sat}} \tag{6.32}$$

Thus, a 741C with $V_{sat} = 13$ V has FPB $= 0.5 \times 10^6/(2\pi 13) = 6.1$ kHz. Exceeding this frequency will yield a distorted as well as reduced output. When applying an amplifier we must make sure that neither its slew-rate limit SR nor its -3-dB frequency f_B is exceeded.

EXAMPLE 6.8. A 741C op amp with ± 15-V supplies is configured as a noninverting amplifier with a gain of 10 V/V. (*a*) If the ac input amplitude is $V_{im} = 0.5$ V, what is the maximum frequency before the output distorts? (*b*) If $f = 10$ kHz, what is the maximum value of V_{im} before the output distorts? (*c*) If $V_{im} = 40$ mV, what is the useful frequency range of operation? (*d*) If $f = 2$ kHz, what is the useful input amplitude range?

Solution.

(*a*) $V_{om} = AV_{im} = 10 \times 0.5 = 5$ V; $f_{max} = \text{SR}/(2\pi V_{om}) = 0.5 \times 10^6/(2\pi 5) \cong 16$ kHz.

(*b*) $V_{om(\max)} = \text{SR}/(2\pi f) = 0.5 \times 10^6/(2\pi 10^4) = 7.96$ V; $V_{im(\max)} = V_{om(\max)}/A = 7.96/10 = 0.796$ V.

(*c*) To avoid slew-rate limiting, keep $f \le 0.5 \times 10^6/(2\pi \times 10 \times 40 \times 10^{-3}) \cong 200$ kHz. Note, however, that $f_B = f_t/A_0 = 10^6/10 = 100$ kHz. The useful range is thus $f \le 100$ kHz, and is dictated by small-signal considerations, rather than slew-rate limiting.

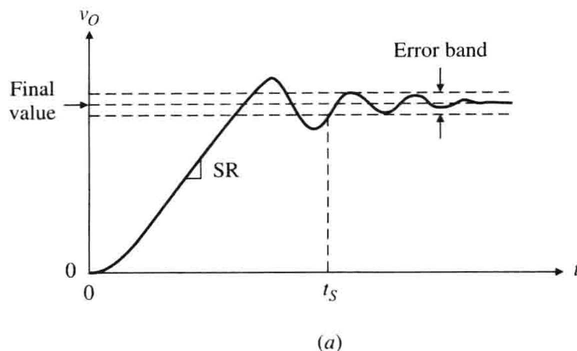
(*d*) $V_{om(\max)} = 0.5 \times 10^6/(2\pi \times 2 \times 10^3) = 39.8$ V. Since this is greater than V_{sat}, or 13 V, the limiting factor is in this case output saturation. Thus, the useful input amplitude range is $V_{im} \le V_{sat}/A = 13/10 = 1.3$ V.

The Settling Time t_S

The rise time t_R and slew rate SR give an indication of how rapidly the output changes, respectively, under small-signal and large-signal conditions. The parameter of greatest concern in many applications is the *settling time* t_S, defined as the time it takes for the response to a large input step to settle and remain within a specified errorband, usually symmetric about its final value. Settling times are typically specified to accuracies of 0.1% and 0.01% of a 10-V input step. As an example, the AD843 op amp has typically $t_S = 135$ ns to 0.01% of a 10-V step.

As shown in Fig. 6.18*a*, t_S is comprised of an initial propagation delay due to higher-order poles, followed by an SR-limited transition to the vicinity of the final value, followed by a period to recover from the overload conditon associated with the SR, and finally settle toward the final equilibrium value. The settling time depends on both linear and nonlinear factors, and is generally a complex phenomenon.[3,4] A fast t_R or a high SR does not necessarily guarantee a fast t_S. For instance, an op amp may settle quickly within 0.1%, but may take considerably longer to settle within 0.01% due to excessively long ringing.

Figure 6.18*b* shows a popular test circuit[5] for the measurement of t_S. The device under test (DUT) is configured as a unity-gain inverting amplifier, while the equal-valued resistors R_3 and R_4 synthesize what is commonly referred to as a *false ground.* Since $v_{FG} = \frac{1}{2}(v_I + v_O)$, with $v_O = -v_I$ we expect $v_{FG} = 0$ V. In practice, because



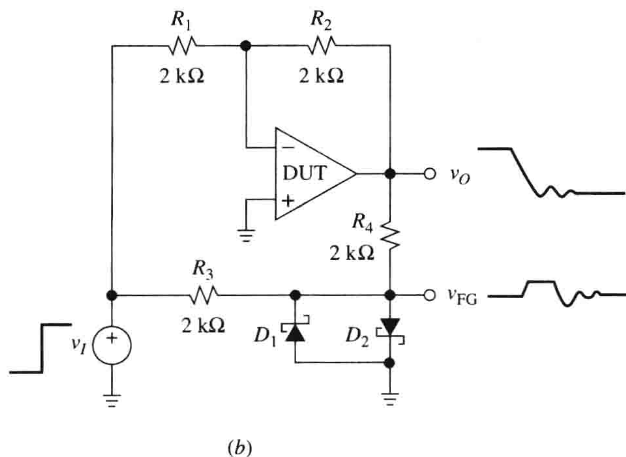


FIGURE 6.18
Settling time t_S, and circuit to measure t_S. (D_1 and D_2 are HP2835 Schottky diodes.)

of the transient due to the op amp, v_{FG} will momentarily deviate from zero and we can observe this deviation to measure t_S. For an error band of $\pm 0.01\%$ of a 10-V step, v_{FG} will have to settle within ± 0.5 mV of its final value. The purpose of the Schottky diodes is to prevent overloading the oscilloscope's input amplifier. To avoid loading by the probe's stray capacitance, v_{FG} can be buffered by means of a JFET source-follower. Consult the data sheets for the recommended test circuit to measure t_S.

In order to fully realize the settling-time capabilities of the op amp, one must pay proper attention to component selection, layout, and grounding; otherwise, the painstaking process of amplifier design can easily be defeated.[5] This includes keeping component leads extremely short, using metal-film resistors, orienting components so as to minimize stray capacitances and connection inductances, properly bypassing the power supplies, and providing separate ground returns for the input, the load, and the feedback network. Fast settling times are particularly desirable in high-speed, high-accuracy D-A converters, sample-and-hold amplifiers, and multiplexed amplifiers.

Slew-Rate Limiting: Causes and Cures

It is instructive to investigate the causes of slew-rate limiting since even a qualitative understanding can better help the user in the op amp selection process. Referring to the block diagram[1] of Fig. 6.19, we observe that as long as the input step amplitude V_m is sufficiently small, the input stage will respond in proportion and yield $i_{O1} = g_{m1}V_m$. By the capacitance law, $dv_O/dt = i_{O1}/C_c = g_{m1}V_m/C_c$, thus confirming that the output rate of change is also propotional to V_m. However, if we overdrive the input stage, i_{O1} will saturate at $\pm I_A$, as depicted in Fig. 5.2*b*. The capacitor C_c will become current-starved, and $(dv_O/dt)_{\max} = I_A/C_c$. This is precisely the slew rate,

$$\text{SR} = \frac{I_A}{C_c} \tag{6.33}$$

Using the 741 op amp working values of Section 5.1, namely, $I_A = 19.6\ \mu\text{A}$ and $C_c = 30$ pF, we estimate SR $= 0.653\ \text{V}/\mu\text{s}$, in reasonable agreement with the data sheets.

It is important to realize that during slew-rate limiting v_N may depart from v_P significantly because of the drastic drop in the open-loop gain brought about by input-stage saturation. During limiting the circuit is insensitive to any high-frequency components at the input. In particular, the virtual-ground condition of the inverting configuration does not hold during limiting. This is confirmed by the shape of v_N in Fig. 6.17.

We can gain additional insight by relating large-signal and small-signal behavior.[1,6] In Eq. (6.8*a*) it was found that $f_t = g_{m1}/(2\pi C_c)$. Solving for C_c and substituting into Eq. (6.33) gives

$$\text{SR} = \frac{2\pi I_A f_t}{g_{m1}} \tag{6.34}$$

This expression points to three different ways of increasing the SR, namely, (*a*) by increasing f_t, (*b*) by reducing g_{m1}, or (*c*) by increasing I_A.

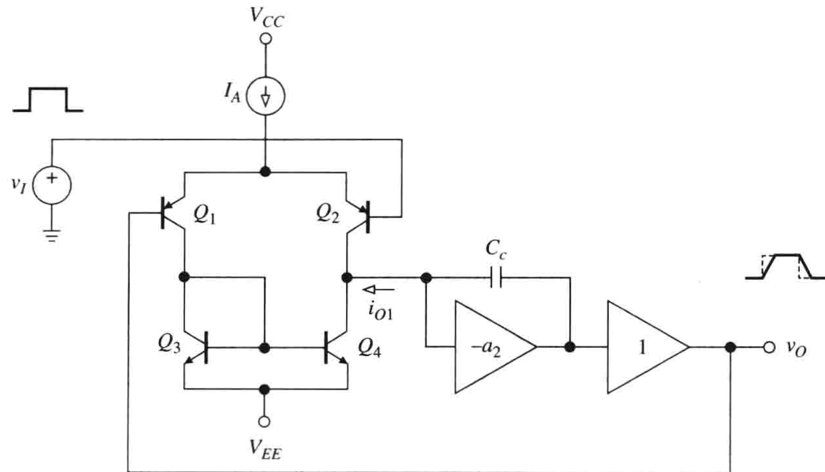


FIGURE 6.19
Op amp model to investigate slew-rate limiting.

In general, an op amp with a high f_t tends to exhibit also a high SR. By Eq. (6.8*a*), f_t can be increased by reducing C_c. This is especially useful in the case of uncompensated op amps, for then the user can specify a compensation network that will also maximize the SR. A popular example is offered by the 301 and 748 op amps, which, when used in high-gain configurations, can be compensated with a smaller C_c value to achieve a higher f_t as well as a higher SR. Even in low-gain applications, frequency-compensation schemes other than the dominant pole are possible, which may improve the SR significantly. Popular examples are the so-called *input-lag* and *feedforward* compensation methods, to be addressed in Chapter 8. For instance, with dominant-pole compensation, the 301 op amp offers dynamic characteristics similar to those of the 741; however, with feedforward compensation, it achieves $f_t = 10$ MHz and SR $= 10$ V/μs.

The second method of increasing the SR is by reducing the input-stage transconductance g_{m1}. For BJT input stages, g_{m1} can be reduced via *emitter degeneration,* which is obtained by including suitable resistances in series with the emitters in the differential input pair to deliberately reduce, or degenerate, transconductance. The LM318 op amp utilizes this technique to achieve SR $= 70$ V/μs with $f_t = 15$ MHz. Alternatively, g_{m1} can be reduced by implementing the differential input pair with FETs, whose transconductance is notoriously lower than that of BJTs for similar biasing conditions. For instance, the TL080 op amp, which is similar to the 741 except for the replacement of the input BJT pair with a JFET pair, offers SR $= 13$ V/μs at $f_t = 3$ MHz. We are now able to appreciate two good reasons for having a JFET input stage: one is to achieve very low input bias and offset currents, and the other is to enhance the slew rate.

The third method of increasing the SR is by increasing I_A. This is especially important in the case of *programmable op amps,* so called because their internal operating currents can be programmed by the user via an external current I_{SET}. (This current is usually set by connecting a suitable external resistor, as specified in the data sheets.) The internal currents, including the quiescent supply current I_Q and the input-stage bias current I_A, are related to I_{SET} in current-mirror fashion, and are thus programmable over a wide range of values. By Eqs. (6.8*b*) and (6.33), both f_t and SR are proportional to I_A, which in turn is proportional to I_{SET}, indicating that the op amp dynamics are also programmable.

6.5 EFFECT OF FINITE GBP ON INTEGRATOR CIRCUITS

Having investigated the frequency response of purely resistive circuits, we now turn to circuits whose feedback networks contain capacitors and thus exhibit frequency-dependent feedback factors. The most popular circuit is the inverting integrator of Fig. 3.6, which serves as a building block not only for filters, but also for signal generators and data converters, to be studied in later chapters. As we know, the ideal inverting-integrator transfer function is

$$H_{\text{ideal}}(jf) = \frac{-1}{jf/f_{0(\text{ideal})}} \qquad (6.35a)$$

Because of the op amp's gain rolloff with frequency, we expect the actual transfer function $H(jf)$ to depart from the ideal. We visualize this departure by means of the PSpice circuit of Fig. 6.20*a*, which uses a 1-MHz GBP op amp with $a_0 = 10^3$ V/V, and has been designed for a *unity-gain frequency* of

$$f_{0(\text{ideal})} = \frac{1}{2\pi RC} = \frac{1}{2\pi 10^5 \times 15.9155 \times 10^{-12}} = 100 \text{ kHz} \tag{6.35b}$$

Figure 6.20*b* indicates that $H(jf)$ approaches $H_{\text{ideal}}(jf)$ only over a limited frequency range, say, from 10^2 to 10^6 Hz, and with a unity-gain frequency (measured with PSpice's cursor) of $f_0 = 89.74$ kHz.

For a more detailed investigation, we start out with the feedback factor, which is easily seen to be a high-pass function of the type

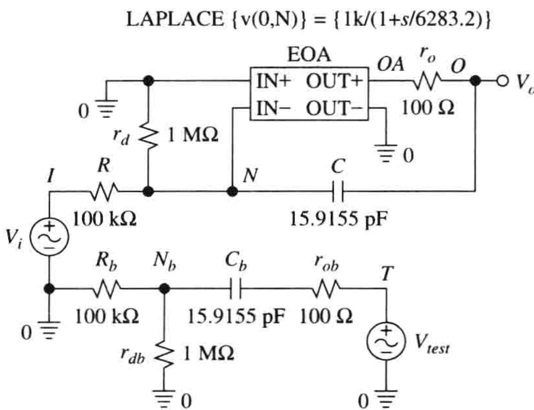
$$\beta(jf) = \beta_\infty \frac{jf/f_1}{1 + jf/f_1} \tag{6.36a}$$

where β_∞ is the high-frequency asymptotic value of β, obtained by letting C become a short circuit, and f_1 is the break frequency formed by C and the equivalent resistance seen by C itself,

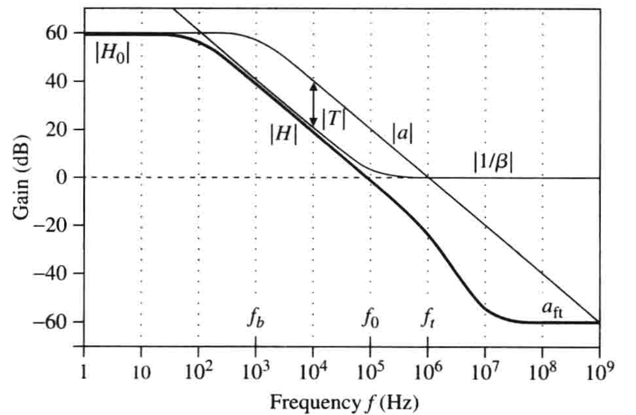
$$\beta_\infty = \frac{R \parallel r_d}{r_o + R \parallel r_d} = 0.9989 \qquad f_1 = \frac{1}{2\pi(R \parallel r_d + r_o)C} = 109.88 \text{ kHz} \tag{6.36b}$$

The loop gain $|T|$, visualized as the *difference* between the $|a|$ and $|1/\beta|$ curves, gives a measure of how close the actual transfer function $H(jf)$ is to ideal. We make the following observations:

1. From f_b to f_0 the loop gain $|T|$ is maximum and approximately frequency-independent.



(*a*)



(*b*)

FIGURE 6.20
(*a*) PSpice integrator and (*b*) its frequency characteristics. The feedback network for plotting $1/\beta$ is repeated at the bottom of (*a*), with its elements identified by subscript *b*. The $|H|$ trace is plotted as `DB(V(O)/V(I))`, the $|a|$ trace as `DB(V(OA)/(-V(N)))`, and the $|1/\beta|$ trace as `DB(V(T)/V(Nb))`.

2. Below f_b, $|T|$ decreases with frequency until it drops to unity at the intercept of the two curves. Below this intercept, C becomes an open circuit, so the op amp will amplify with its full open-loop gain, giving $H_0 = [r_d/(R + r_d)]a_0 = 90.9 \times 10^3 = 909$ V/V = 59.2 dB.
3. Above f_0, $|T|$ decreases again until it drops to unity at the second intercept, in the vicinity of f_t. Above this intercept, the slope of $|H|$ doubles to -40 dB/dec, of which -20 dB/dec are due to the integrator itself, and the additional -20 dB/dec are due to the discrepancy function $D = 1/(1 + 1/T)$.
4. At very high frequencies, C becomes a short circuit, resulting in a feedthrough gain of $a_{\text{ft}} \cong r_o/R = 10^{-3}$ V/V = -60 dB. For its magnitude curve to flatten out, $H(jf)$ must contain a zero pair to neutralize the combined slope due to the pole pair (see Problem 6.43).
5. It is apparent that, at least up to f_t, the integrator behaves like a constant GBP amplifier with GBP $\cong f_0$, a dc gain of H_0, and a -3-dB frequency of about f_0/H_0.

Magnitude and Phase Errors

Ignoring feedthrough because it occurs way above the integrator's useful frequency range, we adapt Eq. (6.13) to the present circuit and approximate

$$H(jf) \cong \frac{-1}{jf/f_{0(\text{ideal})}} \times \frac{1}{1 + 1/T(jf)} \tag{6.37}$$

indicating that $H(jf)$ exhibits both a *magnitude error* and a *phase error*, as per Eq. (6.14). We are particularly interested in the frequency range $f_b \ll f \ll f_1$ over which the loop gain is maximum and

$$T = a\beta \cong \frac{a_0}{jf/f_b} \times \beta_\infty(jf/f_1) = \beta_\infty \frac{f_t}{f_1} \tag{6.38a}$$

Substituting into Eq. (6.37) gives the insightful result

$$H(jf) \cong \frac{-1}{jf/\left[f_{0(\text{ideal})}/(1 + 1/T)\right]} \tag{6.38b}$$

indicating that the effect of the op amp's gain rolloff is to *downshift* the integrator's unity-gain frequency from $f_{0(\text{ideal})}$ to about $f_0 = f_{0(\text{ideal})}/(1+1/T)$. In the present case we have $T = 0.9989 \times 10^6/(109.88 \times 10^3) = 9.09$, so the downshift is from 100 kHz to about $100/(1+1/9.09) = 90.09$ kHz (in fair agreement with the PSpice value of 89.74 kHz).

This downshift is not necessarily bad per se, for we can always *predistort* the value of $f_{0(\text{ideal})}$ so that, by the time it gets divided by $1 + 1/T$, it attains the desired value. In the present circuit, this would require designing for $1/(2\pi RC) = (100\text{ kHz})\times(1+1/9.09) = 111$ kHz, which can be achieved by lowering R from 100 kΩ to $100/111 = 90.09$Ω (88.65 kΩ according to PSpice). In practice, predistortion for low values of T may not be advisable as f_t in Eq. (6.38a) is subject to production variations, drift, and aging. A better bet is to use an op amp with a higher f_t relative to the desired $f_0(\cong f_1)$, since this would raise the value of T and thus reduce the

effect of variations in f_t. A higher f_t will also expand the upper end of the useful frequency range, just like a higher a_0 will expand the lower end of the useful range (see Problem 6.45).

According to Eq. (6.35a), the integrator should provide a phase shift of 90°. In practice, because of the two breakpoints, the shift will depart from 90° at both the lower and upper regions of the useful frequency range. We shall soon see that the latter is a source of concern in integrator-based filters such as dual-integrator loops. To investigate further, let us assume $r_d = \infty$ and $r_o = 0$, so $\beta_\infty = 1$ and the second intercept is *right* at f_t. Under these conditions we can approximate the integrator response in the upper region of the useful frequency range as

$$H(jf) \cong \frac{-1}{jf/f_{0(\text{ideal})}} \frac{1}{1 + jf/f_t} \tag{6.39}$$

indicating a phase error of $\epsilon_\phi = -\tan^{-1}(f/f_t)$. We are particularly interested in ϵ_ϕ in the vicinity of f_0. Since a well-designed integrator has $f_{0(\text{ideal})} \ll f_t$, we can approximate, for $f \ll f_t$,

$$\epsilon_\phi \cong -f/f_t \tag{6.40}$$

We can reduce ϵ_ϕ by introducing a suitable amount of phase lead to counteract the phase lag due to the pole frequency f_t. This process is called *phase-error compensation.*

Passive Compensation of Integrators

The integrator of Fig. 6.21a is compensated by means of an input parallel capacitance C_c. If we specify its value so that $|Z_{C_c}(jf_t)| = R$, or $1/(2\pi f_t C_c) = R$, then the phase lead due to the high-pass action by C_c will compensate for the phase lag due to the low-pass term $1/(1 + jf/f_t)$, thus expanding the frequency range of negligible phase error. This technique, also referred to as *zero-pole cancellation,* requires that

$$C_c = 1/(2\pi R f_t) \tag{6.41}$$

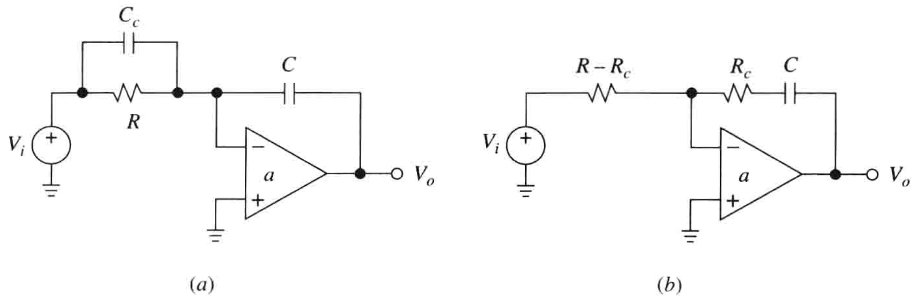


FIGURE 6.21
Passive compensation of the integrator: (a) capacitive, and (b) resistive.

The scheme of Fig. 6.21b achieves a similar result, but by using a feedback series resistance R_c and by decreasing the input resistance from R to $R - R_c$. This method offers better trimming capabilities than capacitive compensation. It can be shown (see Problem 6.50) that letting

$$R_c = 1/(2\pi C f_t) \tag{6.42}$$

will make $H(jf) \to H_{\text{ideal}}$, provided the components are scaled such that the open-loop output impedance z_o is negligible compared to R_c.

Because of manufacturing process variations, the value of f_t is not known precisely, so C_c or R_c must be trimmed for each individual op amp. Even so, compensation is difficult to maintain because f_t is sensitive to temperature and power-supply variations.

Active Compensation of Integrators

The drawbacks of passive compensation are ingeniously avoided with active compensation,[7] so called because it exploits the matching and tracking properties of dual op amps to compensate for the frequency limitations of one device using the very same limitations of the other. Although this technique is general and will be readdressed in Section 8.7, here we focus on compensation of integrators according to the two popular schemes depicted in Fig. 6.22.

With reference to the scheme of Fig. 6.22a, we apply the superposition principle to write

$$V_o = -a_1\left(\frac{1}{1 + jf/f_0}V_i + \frac{jf/f_0}{1 + jf/f_0}A_2 V_o\right) \qquad A_2 = \frac{1}{1 + jf/f_{t2}}$$

where $f_0 = 1/(2\pi RC)$. To find $H = V_o/V_i$, we eliminate A_2, substitute $a_1 \cong f_{t1}/jf$, and let $f_{t2} = f_{t1} = f_t$ to reflect matching. This gives $H(jf) = H_{\text{ideal}} \times 1/(1 + 1/T)$, where

$$\frac{1}{1 + 1/T} = \frac{1 + jf/f_t}{1 + jf/f_t - (f/f_t)^2} = \frac{1 - j(f/f_t)^3}{1 - (f/f_t)^2 + (f/f_t)^4} \tag{6.43}$$

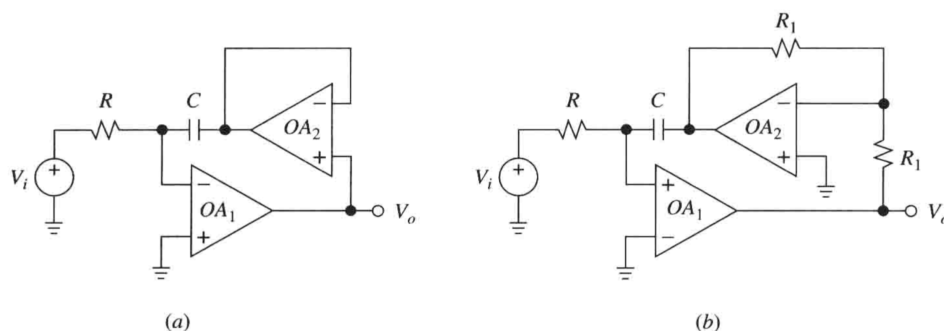


FIGURE 6.22
Active compensation of the integrator: (a) $\epsilon_\phi = -(f/f_t)^3$, and ($b$) $\epsilon_\phi = +f/f_t$.

The last step reveals an interesting property: the rationalization process leads to the mutual cancellation of the first- and second-order terms in f/f_t in the numerator, leaving only the third-order term. We thus approximate, for $f \ll f_t$,

$$\epsilon_\phi \cong -(f/f_t)^3 \tag{6.44}$$

indicating a much smaller error than that of Eq. (6.40). The effect of the compensation scheme of Fig. 6.22*a* is illustrated in Fig. 6.23 by means of PSpice.

Turning next to the scheme of Fig. 6.22*b*, we observe that OA_1 contains the inverting op amp OA_2 in its feedback path, so its input polarities have been interchanged to keep feedback negative. One can prove (see Problem 6.51) that

$$\frac{1}{1+1/T} = \frac{1 + jf/0.5f_t}{1 - jf/f_t - (f/0.5f_t)^2} \cong \frac{1 + jf/f_t}{1 - 3(f/f_t)^2}$$

where we have ignored higher-order terms in f/f_t. We now have

$$\epsilon_\phi \cong +f/f_t \tag{6.45}$$

Though not as small as in Eq. (6.44), this phase error has the advantage of being positive, a feature we shall exploit shortly.

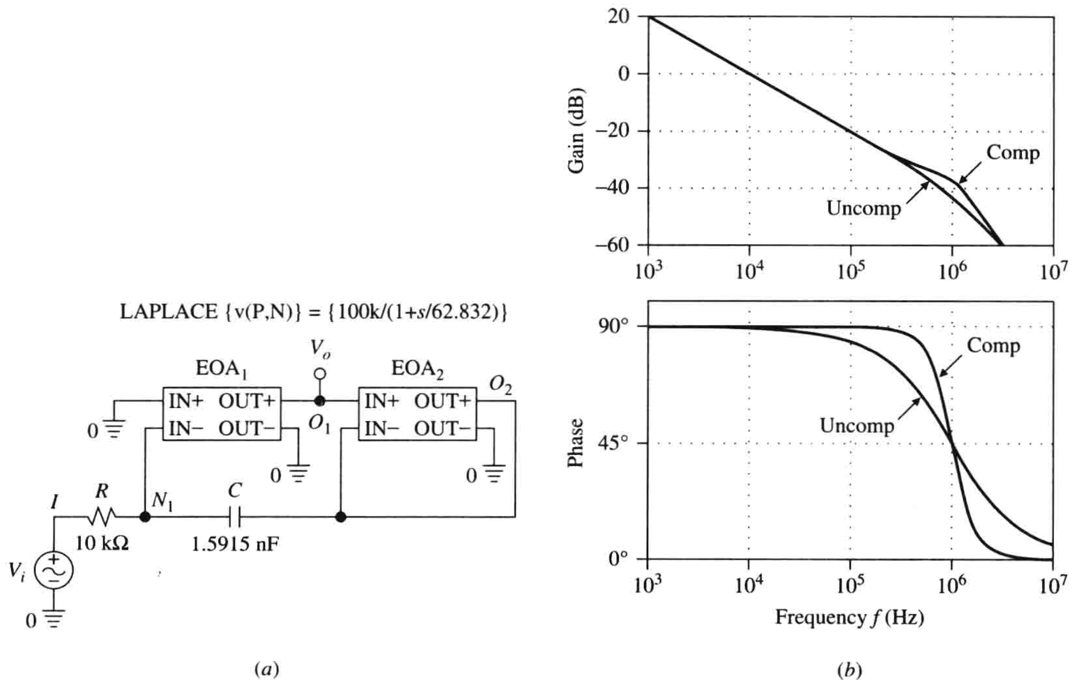


FIGURE 6.23
(*a*) PSpice integrator with active compensation, and (*b*) its Bode plots. For comparison, also shown are the uncompensated plots, obtained with the capacitor's right plate connected directly to V_o.

Q-Enhancement Compensation

It has been found that the effect of nonideal op amps on dual-integrator-loop filters such as the state-variable and biquad varieties is to raise the actual value of Q above the design value predicted under ideal op amp assumptions. This effect, aptly referred to as *Q enhancement,* has been analyzed[8] for the case of the biquad configuration in terms of the phase errors introduced by the two integrators and the third amplifier. The result is

$$Q_{\text{actual}} \cong \frac{Q}{1 - 4Qf_0/f_t} \tag{6.46}$$

where f_0 is the integrator unity-gain frequency, f_t is the op amp transition frequency, and Q is the quality factor in the ideal op amp limit $f_t \to \infty$. As pictured in Fig. 6.24 for a design value of $Q = 25$ and op amps with $f_t = 1$ MHz, Q_{actual} increases with f_0 until it becomes infinite for $f_0 \cong f_t/(4Q) = 10^6/100 = 10$ kHz. At this point the circuit becomes oscillatory.

Besides Q enhancement, the finite GBP of the op amps causes also a shift in the characteristic frequency f_0 of the filter,[9]

$$\frac{\Delta f_0}{f_0} \cong -(f_0/f_t) \tag{6.47}$$

For small Q deviations, Eq. (6.46) gives

$$\frac{\Delta Q}{Q} \cong 4Qf_0/f_t \tag{6.48}$$

Together, these equations indicate the GBP that is needed to contain $\Delta f_0/f_0$ and $\Delta Q/Q$ within specified limits.

EXAMPLE 6.9. Specify suitable components in the biquad filter of Fig. 3.36 to achieve $f_0 = 10$ kHz, $Q = 25$, and $H_{0BP} = 0$ dB, under the constraint that the deviations of f_0 and Q from their design values because of finite GBPs be within 1%.

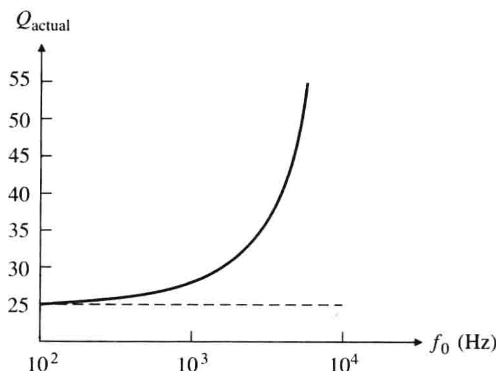


FIGURE 6.24
Q enhancement.

Solution. Use $R_1 = R_2 = R_5 = R_6 = 10\,\text{k}\Omega$, $R_3 = R_4 = 250\,\text{k}\Omega$, and $C_1 = C_2 = 5/\pi$ nF. To meet the f_0 and Q specifications, we need, respectively, $f_t \geq f_0/(\Delta f_0/f_0) = 10^4/0.01 = 1$ MHz, and $f_t \geq 4 \times 25 \times 10^4/0.01 = 100$ MHz. The Q specification is the most demanding, so we need GBP $\geq$ 100 MHz.

The onerous GBP requirements imposed by the Q specification can be relaxed dramatically if we use phase-error compensation to eliminate the Q-enhancement effect. An actual example will better illustrate.

EXAMPLE 6.10. (*a*) Use PSpice to display the band-bass response of the biquad filter of Example 6.9 for the case of op amps with GBP = 1 MHz (use Laplace blocks of the type of Fig. 6.23). (*b*) Calculate the capacitance C_c that, when connected across R_1, will provide passive compensation for all three op amp blocks. Show the compensated response. (*c*) Predistort the values of C_1 and C_2 for a resonant frequency of 10.0 kHz. (*d*) Repeat, but using active compensation. Compare the two circuit implementations, and comment.

Solution.

(*a*) The circuit is shown in Fig. 6.25. The band-pass response without C_c is shown in Fig. 6.26*a* as the "Uncomp" curve. Evidently, the uncompensated circuit suffers from significant Q enhancement.

(*b*) To compensate for the two integrating op amps as well as the unity-gain inverting op amp, whose pole frequency is $f_B = f_t/2$, we use a single capacitance but *four* times as large as that predicted by Eq. (6.41), or $C_c = 4/(2\pi R_1 f_t) = 2/(\pi 10^4 \times 10^6) \cong 64$ pF. The response with C_c in place is shown in Fig. 6.26*a* as the "Comp" curve. Although the effect of Q enhancement has been neutralized, the response still exhibits the frequency shift of Eq. (6.47).

(*c*) To compensate for the frequency shift, we predistort the values of C_1 and C_2 by lowering them by $100 \times f_0/f_t = 1\%$, that is, by lowering them from 1.5915 nF to 1.5756 nF. The ensuing response is shown in Fig. 6.26*b*. Using PSpice's cursor facility, we measure the resonant frequency as 10.0 kHz, and the −3-dB frequencies as $f_L = 9.802$ kHz and $f_H = 10.202$ kHz, thus confirming $Q = 25$.

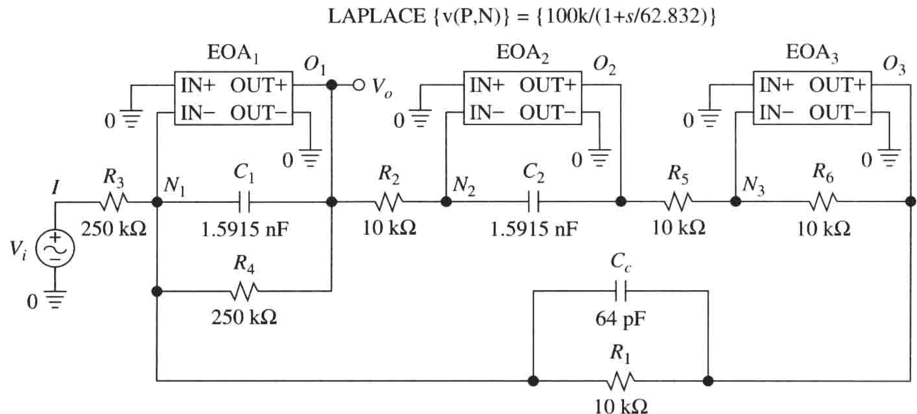


FIGURE 6.25
PSpice realization of the biquad of Example 6.10 using passive compensation.

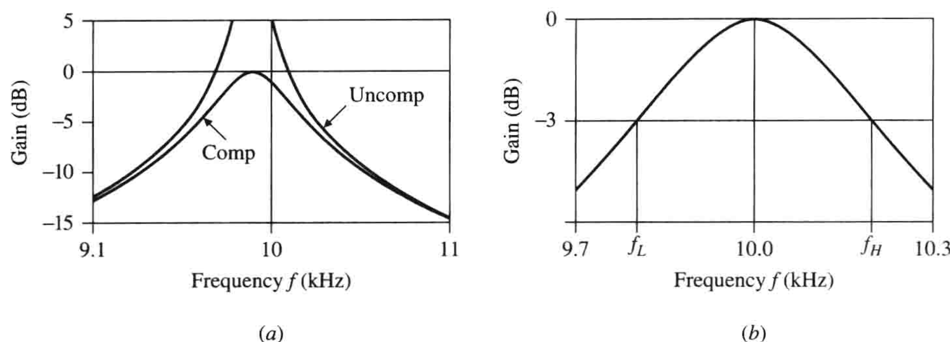


FIGURE 6.26
(a) Band-pass responses of the biquad of Fig. 6.25 with ("Comp") and without ("Uncomp") the compensating capactance C_c in place. (b) The response after predistortion of C_1 and C_2.

(d) For active compensation we rewire the inverting amplifier as in Fig. 6.27. According to Eq. (6.45) the combined phase error of the second integrator and the inverting amplifier is positive and opposite to that of the first integrator, so the overall phase error is neutralized thanks to a mere rewiring of the inverting amplifier! PSpice indicates that there is no frequency shift in this circuit, so C_1 and C_2 are left unchanged. However, there is degradation in the value of Q. Empirically, it is found that raising R_4 to 504 kΩ restores Q to 25.

Before concluding, we wish to point out the above compensation schemes assume single-pole op amps. As seen in Fig. 6.1b for the 741, real-life op amps exhibit additional high-frequency roots whose effect is to further increase the phase error. Consequently, the schemes discussed above must be taken only as starting points, in the sense that a practical circuit may require further tweaking of C_c in the case of passive compensation, or other compensation measures to be discussed in Chapter 8.

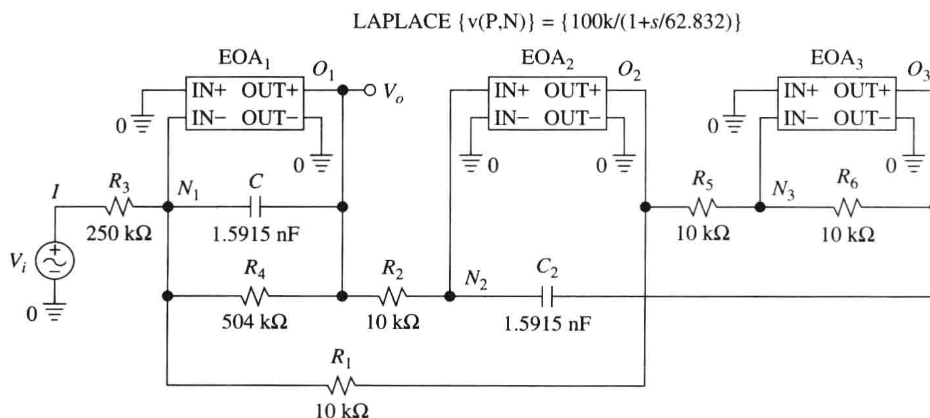


FIGURE 6.27
PSpice realization of the biquad of Example 6.10 using active compensation.

6.6 EFFECT OF FINITE GBP ON FILTERS

In our study of active filters in Chapters 3 and 4, we assumed ideal op amps so we could focus on filter responses alone, without having to worry about op amp idiosyncrasies. We now wish to investigate the effect of op amp gain rolloff with frequency. Unlike circuits with purely resistive feedback, filters exhibit a frequency-dependent feedback factor $\beta(jf)$, so the loop gain $T(jf) = a(jf) \times \beta(jf)| = a(jf)/(1/\beta(jf)$, besides admitting the same poles/zeros as $a(jf)$, will also include the poles/zeros of $\beta(jf)$, which are, respectively, the zeros/poles of $1/\beta(jf)$. As we are about to see, the increased number of roots tends to complicate the calculation of the discrepancy function $D(jf) = 1/[1 + 1/T(jf)]$ and, hence, the transfer function $H(jf)$. Mercifully, when hand analysis becomes prohibitive, we can turn to computer simulation such as SPICE.

First-Order Filters

We start with the low-pass filter of Fig. 3.9*a*, repeated in PSpice form in Fig. 6.28*a* but using the 1-MHz op amp of Fig. 6.3*b*. The circuit has been designed for $H_{0(\text{ideal})} = -R_2/R_1 = -10$ V/V and $f_{0(\text{ideal})} = 1/(2\pi R_2 C) = 20$ kHz. The actual response, shown in Fig. 6.28*b*, approaches the idealized response only up to the intercept of the $|a(jf)|$ and $|1/\beta(jf)|$ curves, in the vicinity of f_t. Cursor measurements give $H_0 = -9.999$ V/V and $f_0 = 16.584$ kHz. Above the intercept, the slope of $|H|$ changes from -20 dB/dec to -40 dB/dec. Further up in frequency, where C acts as a short, we have $|H| \to a_{\text{ft}} \cong r_o/R_1 = 0.01$ V/V ($= -40$ dB).

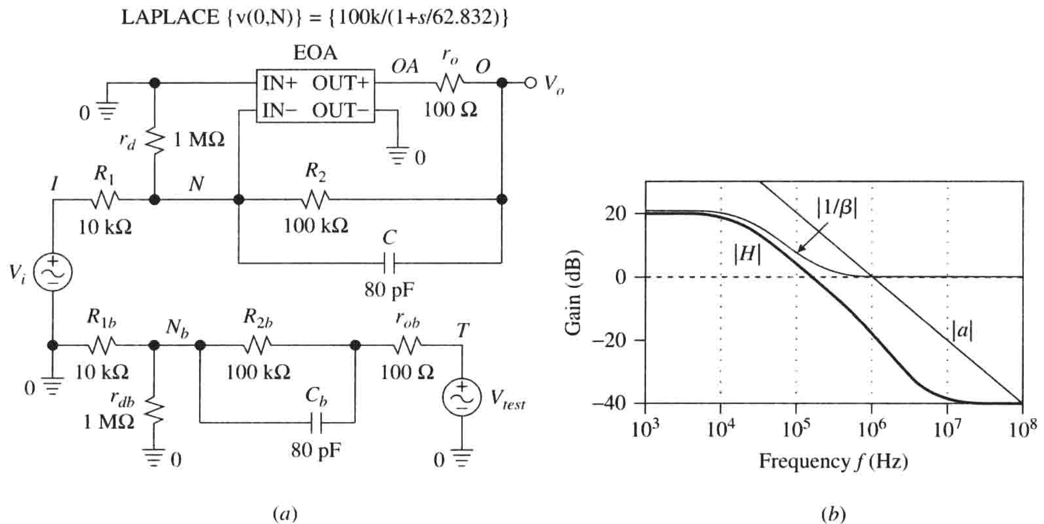


FIGURE 6.28
(*a*) First-order low-pass filter and (*b*) its frequency characteristics.

Adapting Eq. (6.12), we write

$$H(jf) = \frac{V_o}{V_i} \cong \frac{-10}{1 + jf/(20 \times 10^3)} \times \frac{1}{1 + 1/T(jf)} + \frac{0.01}{1 + T(jf)} \qquad (6.49)$$

The open-loop gain $a(jf)$ has a pole frequency at 10 Hz, and the feedback factor $\beta(jf)$ has a zero frequency near 20 kHz and a pole frequency near 200 kHz (see Problem 6.55), so $T(jf)$ has two pole frequencies at 10 Hz and 200 kHz, and a zero frequency at 20 kHz. So long as $|T(jf)| \gg 1$, these roots have a limited impact on $H(jf)$. Not so above the intercept, where $H(jf)$ departs significantly from the ideal, at first because of the pole frequency at the intercept itself, and subsequently because of the zero-frequency pair (in the vicinity of 10 MHz) needed to allow for the high-frequency asymptote $|H(jf)| \rightarrow -40$ dB (not surprisingly, the high-frequency portion of the response is qualitatively similar to that of Fig. 6.20*b* because at high frequencies C dominates over R_2 in Fig. 6.28*a*, making the circuit act as an integrator).

Next, we turn to the high-pass filter of Fig. 3.10*a*, repeated in PSpice form in Fig. 6.29*a* but using the 1-MHz op amp of Fig. 6.3*b*. The circuit has been designed for $H_{0(\text{ideal})} = -R_2/R_1 = -10$ V/V and $f_{0(\text{ideal})} = 1/(2\pi R_1 C) = 1$ kHz. The actual response, shown in Fig. 6.28*b*, approaches the idealized response only up to the intercept of the $|a(jf)|$ and $|1/\beta(jf)|$ curves, in the vicinity of 100 kHz, with measured values of $H_0 = -9.987$ V/V and $f_0 = 990$ Hz. Above the intercept the response changes to a low-pass type. Since feedthrough is more than a decade lower than in Fig. 6.28*a*, we ignore it in the present case and adapt Eq. (6.12) to write

$$H(jf) = \frac{V_o}{V_i} \cong -10 \frac{f/10^3}{1 + jf/10^3} \times \frac{1}{1 + 1/T(jf)} \qquad (6.50)$$

Presently, $\beta(jf)$ has a pole frequency near 100 Hz and a zero frequency at 1 kHz (see Problem 6.57), so $T(jf)$ has two pole frequencies at 10 Hz and 100 Hz, and a zero frequency at 1 kHz. So long as $|T(jf)| \gg 1$, these roots have little impact on

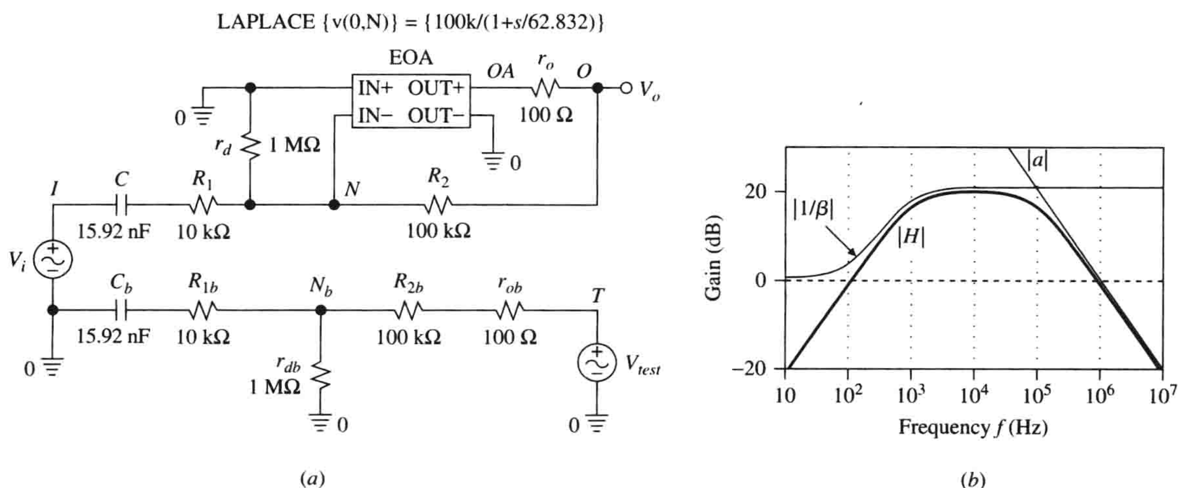


FIGURE 6.29
(*a*) First-order high-pass circuit and (*b*) its frequency response.

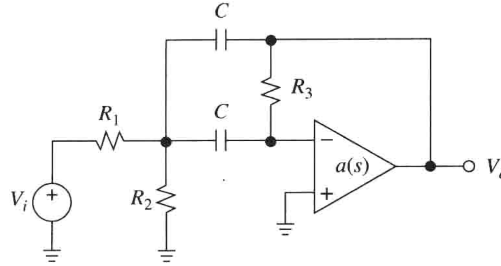


FIGURE 6.30
The multiple-feedback band-pass filter.

$H(jf)$. Not so above the intercept, due to the pole frequency at the intercept itself. In fact, the overall response is that of a wideband band-pass filter (see Problem 6.58).

Second-Order Filters

We now turn to the multiple-feedback configuration of Fig. 6.30 as a popular second-order filter representative. It can be shown (see Exercise 6.1) that for the case of an op amp with gain $a(s)$, $r_d = \infty$, and $r_o = 0$, the circuit admits the transfer function

$$H(s) = H_{0BP} \frac{(s/\omega_0)/Q}{\dfrac{s^2}{\omega_0^2} + \dfrac{1}{Q}\dfrac{s}{\omega_0} + 1 + \dfrac{1}{a(s)}\left(\dfrac{s^2}{\omega_0^2} + \dfrac{2Q^2+1}{Q}\dfrac{s}{\omega_0} + 1\right)} \tag{6.51}$$

where H_{0BP}, ω_0, and Q are as in Eq. (3.71). It is apparent that as long as $a(s)$ is very large, the effect of the denominator block within parentheses will be negligible, resulting in $H(s) \to H_{0BP}H_{BP}(s)$. But, as $a(s)$ rolls off with frequency, the block within parentheses becomes increasingly relevant, leading to a shift in all three filter parameters. Moreover, the additional pole frequency arising at f_t will make $H(s)$ roll off with an ultimate slope of -40 dB/dec instead of -20 dB/dec.

EXERCISE 6.1. Derive Eq. (6.51).

We are primarily interested in the deviations of the resonance frequency and the -3-dB bandwidth from their design values. It has been shown[9] that as long as $Qf_0 \ll f_t$, we have

$$\frac{\Delta f_0}{f_0} \cong -\frac{\Delta Q}{Q} \cong -Qf_0/f_t \tag{6.52}$$

Evidently the product $Q \times f_0$ provides an indication of how demanding the filter specifications are in terms of the GBP.

EXAMPLE 6.11. Using 10-nF capacitances, specify suitable components in the circuit of Fig. 6.30 for $H_{0BP} = 0$ dB, $f_0 = 10$ kHz, $Q = 10$, and a BW deviation from its design value due to finite GBP of 1% or less.

Solution. Using Eqs. (3.72) and (3.73), we get $R_1 = 15.92$ kΩ, $R_2 = 79.98$ Ω, and $R_3 = 31.83$ kΩ. Since BW $= f_0/Q$, Eq. (6.52) gives $\Delta\text{BW}/\text{BW} \cong -2Qf_0/f_t$. Consequently, GBP $\geq 2 \times 10 \times 10^4/0.01 = 20$ MHz.

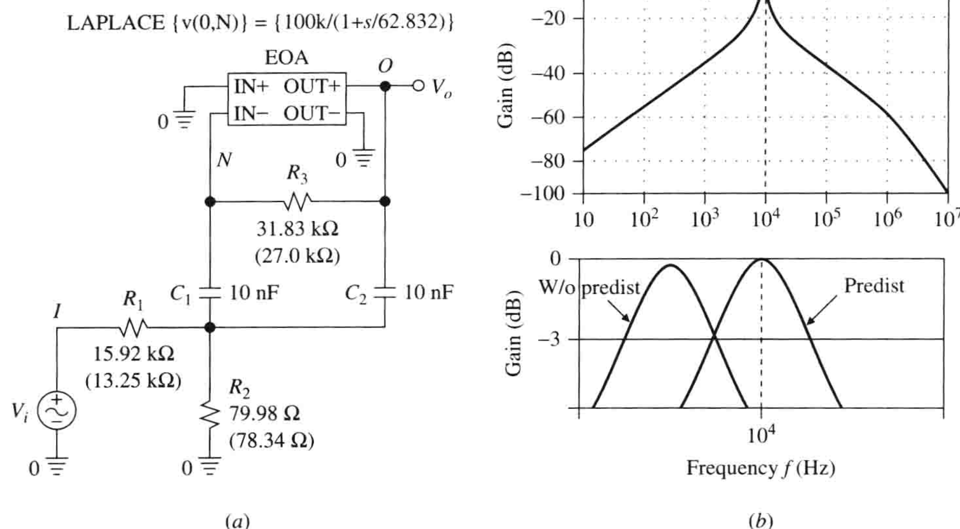


FIGURE 6.31
(a) Multiple-feedback band-pass filter (predistorted values shown inside parentheses), and (b) expanded view of the frequency response before predistortion (top) and with and without predistortion (bottom).

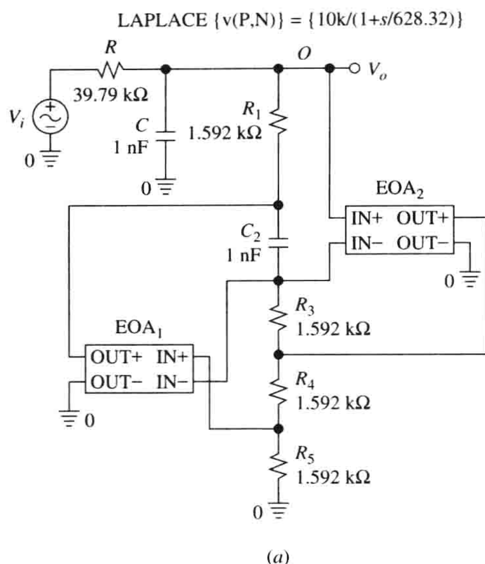
An alternative to using high-GBP op amps is to predistort the filter parameters so as to make the actual values coincide with those given in the specifications. In this respect, PSpice simulation provides an invaluable tool in determining the amount of predistortion required for a given value of f_t.

EXAMPLE 6.12. Design a filter meeting the specifications of Example 6.11 with a 1-MHz op amp.

Solution. With $f_t = 1$ MHz we get $Qf_0/f_t = 0.1$, so by Eq. (6.52) we expect a decrease in f_0 and an increase in Q on the order of 10%. For more accurate estimates we use the PSpice circuit of Fig. 6.31a, whose overall response, shown at the top of Fig. 6.31b, confirms a frequency downshift as well as a high-frequency rolloff of −40 dB/dec instead of −20 dB/dec. Cursor measurements on the expanded plot shown at the bottom give $H_{0BP} = 0.981$ V/V $= -0.166$ dB, $f_L = 8.73$ kHz, $f_H = 9.55$ kHz, and $f_0 = 9.13$ kHz, so $Q = 9.13/(9.55 - 8.73) = 11.3$.

To achieve the desired parameter values, we redesign the circuit for the predistorted values of $f_0 = 10(10/9.13) = 10.95$ kHz, $Q = 10(10/11.3) = 8.85$, and $H_{0BP} = 1/0.981 = 1.02$ V/V. Using again Eqs. (3.72) and (3.73), we find the resistance values shown within parentheses in Fig. 6.31a. Cursor measurements on the predistorted response at the bottom of Fig. 6.31b give $f_0 = 10.0$ kHz and $Q = 10$.

EXAMPLE 6.13. Use PSpice to investigate the effect of implementing the DABP filter of Example 4.8 with op amps having $a_0 = 80$ dB and GBP = 1 MHz. Discuss your findings.



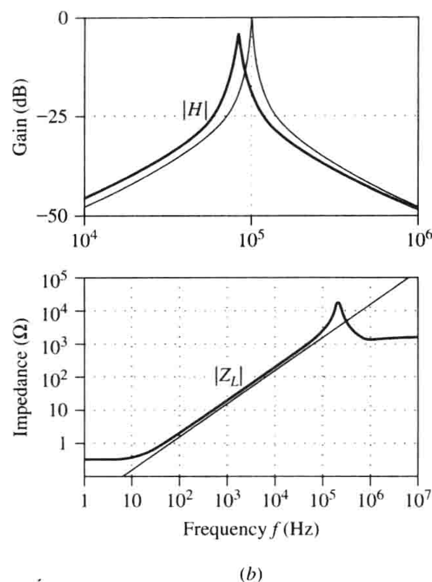


FIGURE 6.32
(*a*) DABP filter of Example 6.13. (*b*) Frequency response $|H|$ (top) and synthetic inductor impedance $|Z_L|$ (bottom). (Thin traces show the responses for the case of ideal op amps.)

Solution. Using the PSpice circuit of Fig. 6.32*a*, we get the response of Fig. 6.32*b* (top), indicating that the effect of nonideal op amps is a decrease in f_0, Q, and H_0. To gain better insight, it is convenient to plot also the impedance Z_L of the synthetic inductor, obtained by removing R and C, subjecting the synthetic inductor circuitry to a test currrent I_t, and then letting $Z_L = V_o/I_t$. It is apparent that the effect of the gain rolloff of the op amps is to raise the effective value of L, causing it even to peak between 10^5 and 10^6 Hz, and then to approach the high-frequency asymptote $Z_L \to R_1$. Fortunately, above 10^5 Hz, C takes over to ensure the eventual −20-dB rolloff of $|H|$. We can significantly improve the response by using faster op amps. For instance, rerunning PSpice after raising their GBP from 1 MHz to 10 MHz, we get a response much closer to ideal.

Concluding Remarks

Looking back at the integrator as well as the filter examples of the current section, we conclude that the open-loop gain rolloff with frequency tends to cause a downshift in the frequency characteristics of the filter as well as a steeper rate of descent at high frequencies, at least up to the point at which feedthrough takes over. The interested reader is referred to the literature[9] for detailed studies of the effect of finite GBP on filters. Within the scope of this book, we limit ourself to finding the actual response via computer simulation, using the more realistic SPICE macromodels provided by the manufacturers, and then applying predistortion in the manner of Examples 6.10 and 6.12. As rule of thumb, one should select an op amp with a GBP at least an order of magnitude higher than the filter product Qf_0 in order to reduce the effect of GBP variations due to environmental and manufacturing process variations.

6.7 CURRENT-FEEDBACK AMPLIFIERS[10]

The op amps considered so far are also referred to as *voltage-feedback amplifiers* (VFAs) because they respond to voltages. As we know, their dynamics are limited by the gain-bandwidth product and the slew rate. By contrast, *current-feedback amplifiers* (CFAs) exploit a circuit topology that emphasizes current-mode operation, which is inherently much faster than voltage-mode operation because it is less prone to the effect of stray node-capacitances. Fabricated using high-speed complementary bipolar processes, CFAs can be orders of magnitude faster than VFAs.

As shown in the simplified diagram of Fig. 6.33, a CFA consists of three stages: (*a*) a *unity-gain input buffer,* (*b*) a pair of *current mirrors,* and (*c*) an *output buffer.* The input buffer is based on the push-pull pair Q_1 and Q_2, whose purpose is to provide very low impedance at its output node v_N, which also acts as the inverting input of the CFA. In the presence of an external network, the push-pull pair can easily source or sink a substantial current i_N, though we shall see that in steady state i_N approaches zero. Q_1 and Q_2 are driven by the emitter followers Q_3 and Q_4, whose purpose is to raise the impedance and lower the bias current at the noninverting input v_P. The followers also provide suitable *pn*-junction voltage drops to bias Q_1 and Q_2 in the forward-active region and thus reduce crossover distortion. By design, the input buffer forces v_N to track v_P. This is similar to ordinary VFAs, except that the latter force v_N to track v_P via negative feedback.

Any current drawn at node v_N by the external network causes an imbalance between the currents of the push-pull pair,

$$i_1 - i_2 = i_N \tag{6.53}$$

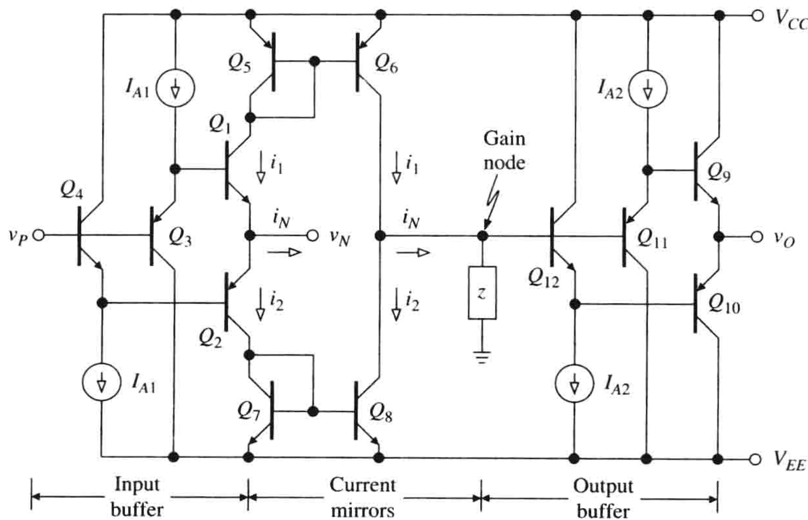


FIGURE 6.33
Simplified circuit diagram of a current-feedback amplifier.

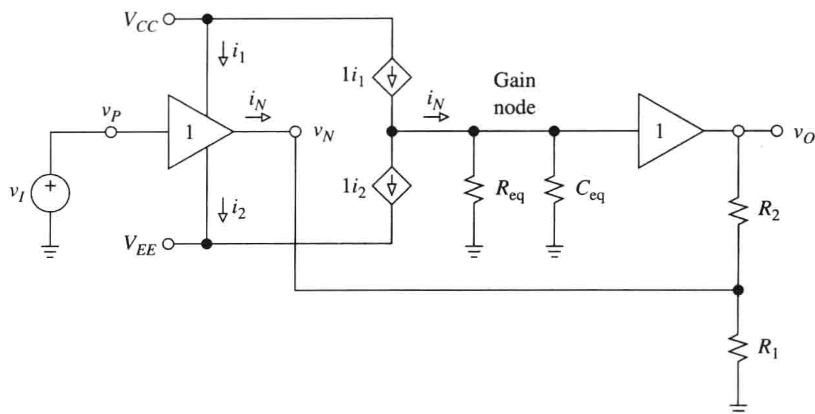


FIGURE 6.34
Block diagram of a CFA configured as a noninverting amplifier.

The current mirrors Q_5-Q_6 and Q_7-Q_8 replicate i_1 and i_2 and sum them at a common node called the *gain node.* The voltage of this node is buffered to the outside by another unity-gain buffer made up of Q_9 through Q_{12}. Ignoring the input bias current of this buffer, we can write, by Ohm's law,

$$V_o = z(jf)I_n \tag{6.54}$$

where $z(jf)$, the net equivalent impedance of the gain node toward ground, is called the *open-loop transimpedance gain.* This transfer characteristic is similar to that of a VFA, except that the error signal i_N is a current rather than a voltage, and the gain $z(jf)$ is in volts per ampere rather than volts per volt. For this reason CFAs are also called *transimpedance amplifiers.*

The relevant CFA features are summarized in the block diagram of Fig. 6.34, where z has been decomposed into the *transresistance* component R_{eq} and *transcapacitance* component C_{eq}. Letting $z(jf) = R_{eq} \parallel [1/(j2\pi f C_{eq})]$ and expanding, we get

$$z(jf) = \frac{z_0}{1 + jf/f_b} \tag{6.55}$$

$$f_b = \frac{1}{2\pi R_{eq} C_{eq}} \tag{6.56}$$

where $z_0 = R_{eq}$ is the dc value of $z(jf)$. The gain $z(jf)$ is approximately constant from dc to f_b; thereafter it rolls off with frequency at the rate of -1 dec/dec. Typically, R_{eq} is on the order of 10^6 Ω (which makes z_0 on the order of 1 V/μA), C_{eq} on the order of 10^{-12} F, and f_b on the order of 10^5 Hz.

EXAMPLE 6.14. The CLC401 CFA has $z_0 \cong 0.71$ V/μA and $f_b \cong 350$ kHz. (*a*) Find C_{eq}. (*b*) Find i_N for $v_O = 5$ V (dc).

Solution.

(*a*) $R_{eq} \cong 710$ kΩ, so $C_{eq} = 1/(2\pi R_{eq} f_b) \cong 0.64$ pF.
(*b*) $i_N = v_O/R_{eq} \cong 7.04$ μA.

Closed-Loop Gain

Figure 6.35*a* shows a simplified CFA model, along with a negative-feedback network. Whenever an external signal V_i tries to unbalance the CFA inputs, the input buffer begins sourcing (or sinking) an imbalance current I_n. By Eq. (6.54), this current causes V_o to swing in the positive (or negative) direction until the original imbalance is neutralized via the negative-feedback loop, thus confirming the role of I_n as error signal.

Applying the superposition principle, we can write

$$I_n = \frac{V_i}{R_1 \parallel R_2} - \frac{V_o}{R_2} \tag{6.57}$$

Clearly, the feedback signal V_o/R_2 is a current, and the feedback factor $\beta = 1/R_2$ is now in amperes per volt. Substituting into Eq. (6.54) and collecting gives the *closed-loop gain*

$$A(jf) = \frac{V_o}{V_i} = \left(1 + \frac{R_2}{R_1}\right)\frac{1}{1 + 1/T(jf)} \tag{6.58}$$

$$T(jf) = \frac{z(jf)}{R_2} \tag{6.59}$$

where $T(jf)$ is called the *loop gain.* This name stems from the fact that a current flowing around the loop is first multiplied by $z(jf)$ to be converted to a voltage, and then divided by R_2 to be converted back to a current, thus experiencing an overall gain of $T(jf) = z(jf)/R_2$. In the decade plot of $|z|$ and $|1/\beta|$ of Fig. 6.35*b* we can visualize the decade value of $|T|$ as the decade difference between the two curves. For instance, if at a given frequency $|z| = 10^5$ V/A and $|1/\beta| = 10^3$ V/A, then $|T| = 10^{5-3} = 10^2$.

In their effort to maximize $T(jf)$ and thus reduce the gain error, manufacturers strive to maximize $z(jf)$ relative to R_2. Consequently, the inverting-input current $I_n = V_o/z$ will be very small, even though this is the low-impedance output node

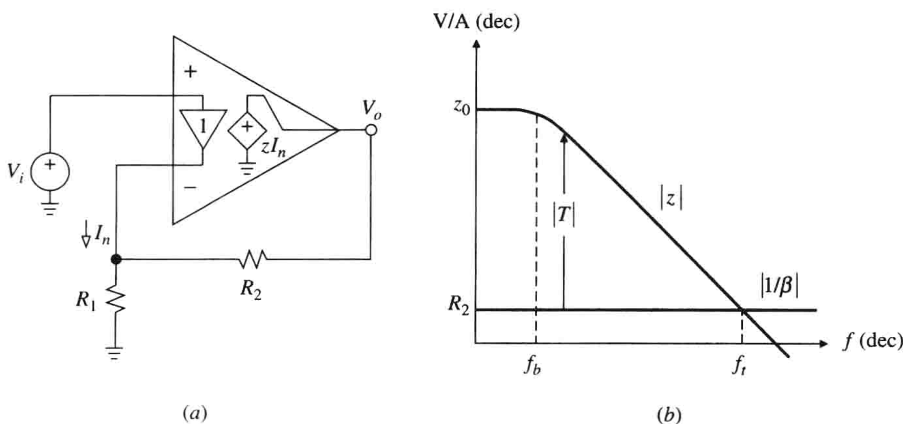


FIGURE 6.35
Noninverting CFA amplifier, and graphical method to visualize the loop gain $|T|$.

of a buffer. In the limit $z \to \infty$ we obtain $I_n \to 0$, indicating that a *CFA will ideally provide whatever output is needed to drive I_n to zero.* Thus, the *input voltage constraint*

$$V_n \to V_p \tag{6.60a}$$

and the *input current constraints*

$$I_p \to 0 \qquad I_n \to 0 \tag{6.60b}$$

hold also for CFAs, though for different reasons than for VFAs. Equation (6.60*a*) holds by design in a CFA, and by negative-feedback action in a VFA; Eq. (6.60*b*) holds by negative-feedback action in a CFA, and by design in a VFA. We can apply these constraints to the analysis of CFA circuits, very much like in the analysis of conventional VFAs.[11]

CFA Dynamics

To investigate the dynamics of the CFA of Fig. 6.34, we substitute Eq. (6.55) into Eq. (6.59), and then into Eq. (6.58). This gives, for $z_0/R_2 \gg 1$,

$$A(jf) = A_0 \times \frac{1}{1 + jf/f_t} \tag{6.61}$$

$$A_0 = 1 + \frac{R_2}{R_1} \qquad f_t = \frac{1}{2\pi R_2 C_{eq}} \tag{6.62}$$

where A_0 and f_t are, respectively, the *closed-loop dc gain* and *bandwidth.* With R_2 in the kiloohm range and C_{eq} in the picofarad range, f_t is typically in the range of 10^8 Hz. We observe that for a given CFA, the closed-loop bandwidth depends on only R_2. We can thus use R_2 to set f_t, and then adjust R_1 to set A_0. The ability to control gain independently of bandwidth constitutes the first major advantage of CFAs over conventional op amps. Bandwidth constancy is illustrated in Fig. 6.36*a*.

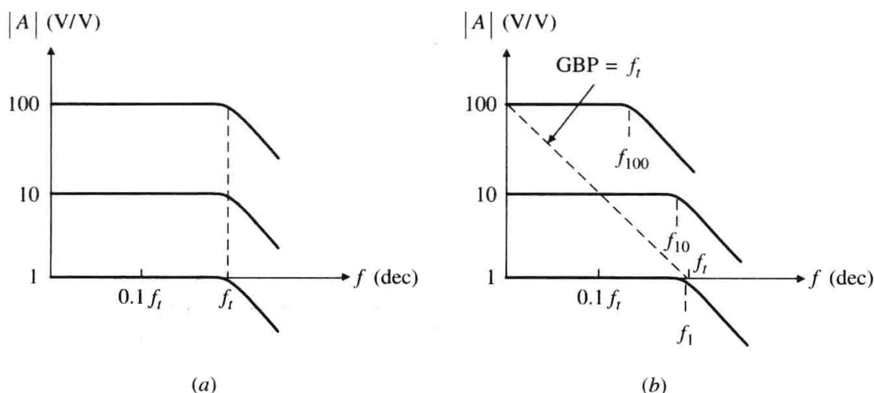


FIGURE 6.36
Closed-loop bandwidth as a function of gain for (*a*) an ideal CFA and (*b*) a practical CFA.

Next, we investigate the transient response. Applying a step $v_I = V_{im}u(t)$ to the circuit of Fig. 6.35*a* will, by Eq. (6.57), result in the current $i_N = V_{im}/(R_1 \parallel R_2) - v_O/R_2$. With reference to Fig. 6.34, we can also write $i_N = v_O/R_{\text{eq}} + C_{\text{eq}}dv_O/dt$. Eliminating i_N, we get, for $R_2 \ll R_{\text{eq}}$,

$$R_2 C_{\text{eq}} \frac{dv_O}{dt} + v_O = A_0 V_{im}$$

whose solution is $v_O = A_0 V_{im}[1 - \exp(t/\tau)]u(t)$,

$$\tau = R_2 C_{\text{eq}} \tag{6.63}$$

The response is an exponential transient regardless of the input step magnitude, and the time constant governing it is set by R_2 regardless of A_0. For instance, a CLC401 op amp with $R_2 = 1.5\ \text{k}\Omega$ has $\tau = 1.5 \times 10^3 \times 0.64 \times 10^{-12} \cong 1$ ns. The rise time is $t_R = 2.2\tau \cong 2.2$ ns, and the settling time within 0.1% of the final value is $t_S \cong 7\tau \cong 7$ ns, in reasonable agreement with the data-sheet values $t_R = 2.5$ ns and $t_S = 10$ ns.

Since R_2 controls the closed-loop dynamics, data sheets usually recommend an optimum value, typically in the range of $10^3\ \Omega$. For voltage follower operation R_1 is removed, but R_2 must be left in place to set the dynamics of the device.

Higher-Order Effects

According to the above analysis, once R_2 has been set, the dynamics appear to be unaffected by the closed-loop gain setting. However, the bandwidth and rise time of a practical CFA do vary with A_0 somewhat, though not as drastically as in VFAs. The main reason is the nonzero output resistance r_n of the input buffer, whose effect is to reduce the loop gain somewhat, degrading the closed-loop dynamics in proportion. Using the more realistic CFA model of Fig. 6.37*a* we get, by the superposition principle, $I_n = V_i/[r_n + (R_1 \parallel R_2)] - \beta V_o$, where the feedback factor β is found

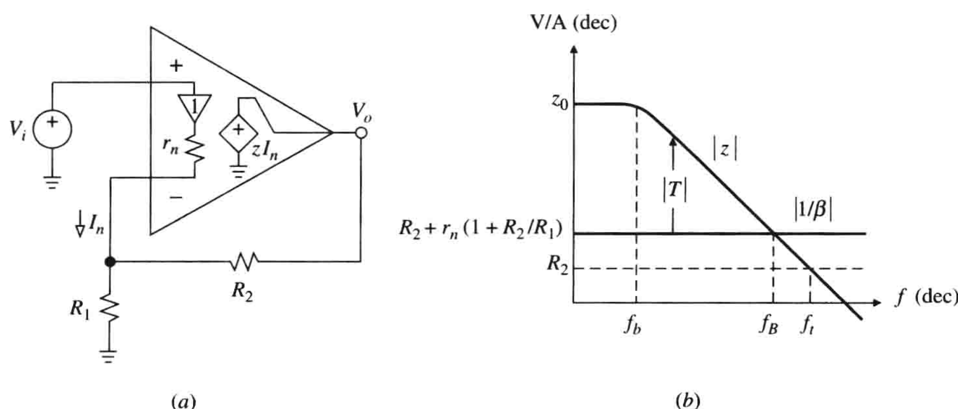


FIGURE 6.37
Effect of the output impedance r_n of the input buffer.

using the current-divider formula and Ohm's law,

$$\beta = \frac{R_1}{R_1 + r_n} \times \frac{1}{R_2 + (r_n \parallel R_1)} = \frac{1}{R_2 + r_n(1 + R_2/R_1)} \tag{6.64}$$

Clearly, the effect of r_n is to shift the $|1/\beta|$ curve upward, from R_2 to $R_2 + r_n(1 + R_2/R_1)$. As pictured in Fig. 6.37*b*, this causes a decrease in the crossover frequency, which we shall now denote as f_B. This frequency is obtained by letting $f_t \to f_B$ and $R_2 \to R_2 + r_n(1 + R_2/R_1)$ in Eq. (6.62). The result can be put in the form

$$f_B = \frac{f_t}{1 + r_n/(R_1 \parallel R_2)} \tag{6.65}$$

where now f_t is the extrapolated value of f_B in the limit $r_n \to 0$.

EXAMPLE 6.15. A certain CFA has $f_t = 100$ MHz for $1/\beta = 1.5$ V/mA. If $R_2 =$ 1.5 kΩ and $r_n = 50$ Ω, find R_1, f_B, and t_R for $A_0 = 1$ V/V, 10 V/V, and 100 V/V. Comment on your results.

Solution. By Eqs. (6.62) and (6.65), we can write, for the present circuit,

$$R_1 = R_2/(A_0 - 1)$$
$$f_B = 10^8/(1 + A_0/30)$$

Moreover, $t_R \cong 2.2/2\pi f_B$. For $A_0 = 1$, 10, and 100 V/V we get, respectively, $R_1 = \infty$, 166.7 Ω, and 15.15 Ω; $f_B = 96.8$ MHz, 75.0 MHz, and 23.1 MHz; $t_R = 2.2/(2\pi \times 96.8 \times 10^6) = 3.6$ ns, 4.7 ns, and 15.2 ns. The bandwidth reductions, depicted in Fig. 6.36*b*, still compare favorably with those of a VFA, whose bandwidth would be reduced, respectively, by 1, 10, and 100.

The values of R_1 and R_2 can be predistorted to compensate for bandwidth reduction. We first find R_2 for a given f_B at a given A_0; then we find R_1 for the given A_0.

EXAMPLE 6.16. (*a*) Redesign the amplifier of Example 6.15 so that with $A_0 = 10$ V/V it has $f_B = 100$ MHz rather than 75 MHz. (*b*) Assuming $z_0 = 0.75$ V/μA, find the dc gain error.

Solution.

(*a*) For $f_B = 100$ MHz we need $R_2 + r_n(1 + R_2/R_1) = 1.5$ V/mA, or $R_2 = 1500 - 50 \times 10 = 1$ kΩ. Then, $R_1 = R_2/(A_0 - 1) = 10^3/(10 - 1) = 111$ Ω.
(*b*) $T_0 = \beta z_0 = (1/1500)0.75 \times 10^6 = 500$. The dc gain error is $\epsilon \cong -100/T_0 = -0.2\%$.

It is apparent that the presence of r_n tends to degrade CFA dynamics. More recent CFA architectures exploit internal negative feedback around the input buffer to reduce its effective output resistance significantly. One example is the OPA684 CFA (look up its data sheets online), which, with an effective inverting-input resistance of 2.5 Ω, allows for the gain-setting element to be set with considerable freedom from bandwidth interaction.

Applying CFAs

Though we have focused on the noninverting amplifier, we can configure a CFA for other familiar topologies.[11] For instance, if we lift R_1 off ground in Fig. 6.35*a*, and apply V_i via R_1 with the noninverting input at ground, we obtain the familiar

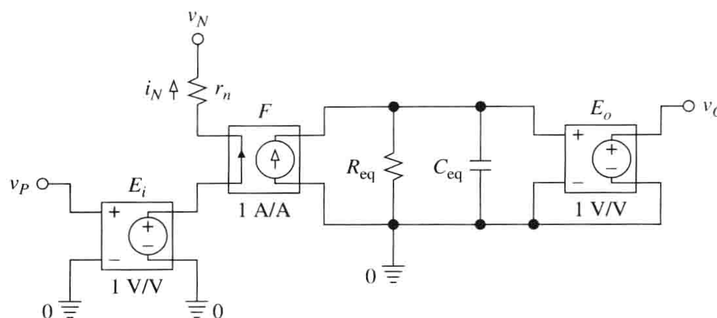


FIGURE 6.38
Simplified PSpice model of a CFA, using VCVSs E_i and E_o, and CCCS F.

inverting amplifier. Its dc gain is $A_0 = -R_2/R_1$, and its bandwidth is given by Eq. (6.65). Likewise, we can configure CFAs as summing or difference amplifiers, *I-V* converters, and so forth. Except for its much faster dynamics, a CFA works much like a VFA, but with one notorious exception that will be explained in Chapter 8: it must never include a direct capacitance between its output and inverting-input pins, since this tends to make the circuit oscillatory. In fact, stable amplifier operation requires that $1/\beta \geq (1/\beta)_{\min}$, where $(1/\beta)_{\min}$ is also given in the data sheets.

Compared to VFAs, CFAs generally suffer from poorer input-offset-voltage and input-bias-current characteristics. Moreover, they afford lower dc loop gains, usually on the order of 10^3 or less. Finally, having much wider bandwidths, they tend to be noisier. CFAs are suited to moderately accurate but very high-speed applications.

PSpice Models

CFA manufacturers provide macromodels to facilitate the application of their products. Alternatively, the user can create simplified models for a quick test of such characteristics as noise and stability. Figure 6.38 shows one such model.

EXAMPLE 6.17. Use PSpice to display the (*a*) frequency and (*b*) transient responses of the amplifier of Example 6.15. For the frequency response show the magnitude plots of $z(jf)$, $1/\beta$, and $A(jf)$, and measure the dc gain error. For the transient response show $v_O(t)$ and $i_N(t)$ if $v_I(t)$ is a 1-V step. Comment on the various plots.

Solution.

(*a*) Using the circuit of Fig. 6.39, we generate the frequency plots of Fig. 6.40*a*. Measuring the dc gain with the cursor, we get $A_0 = 9.981$ V/V, indicating a dc error of $100(9.981 - 10)/10 = -0.19\%$, in agreement with Example 6.15*b*.

(*b*) Next, we change the ac input source V_i of Fig. 6.39 to a 1-V step and perform the transient analysis, whose results are depicted in Fig. 6.40*b*. Since the crossover frequency in Fig. 6.40*a* is 100 MHz, the time constant governing the transient is $\tau = 1/(2\pi 10^8) = 1.59$ ns. Note also the current spike through r_n, which starts at the fairly high value $i_N(0) = v_I(0)/(r_n + R_1 \| R_2) = 1/(50 + 111.1\|1,000) = 6.7$ mA, but as the transient dies out, it drops to the much smaller value $i_N(\infty) = v_O(\infty)/R_{eq} \cong 10/(750\ \text{k}\Omega) = 13.3\ \mu$A.

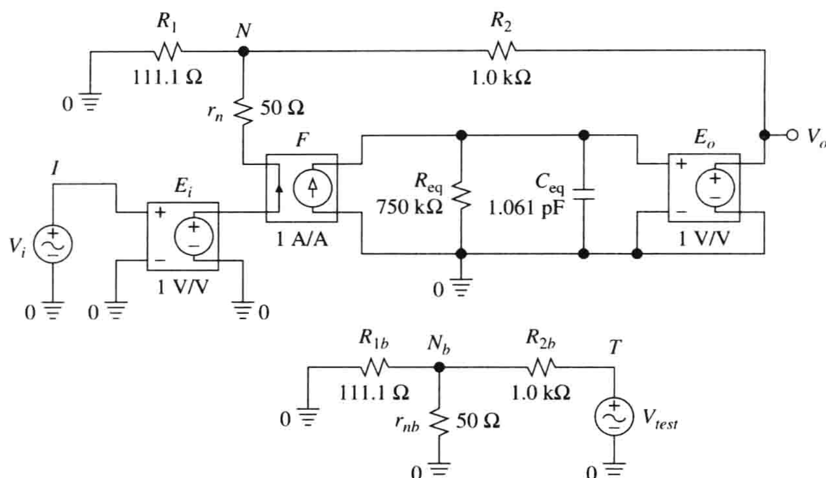


FIGURE 6.39
PSpice circuit to plot the frequency and transient responses of the CFA circuit of Example 6.15. The $|z|$ trace is plotted on a logarithmic scale as `V(O)/(-I(rn))`, the $|1/\beta|$ trace as `V(T)/I(rnb)`, and the $|A|$ trace as `DB(V(O)/V(I))`.

High-Speed Voltage-Feedback Amplifiers

The availability of high-speed complementary bipolar processes and the emergence of applications requiring increased speeds have led to the development of faster

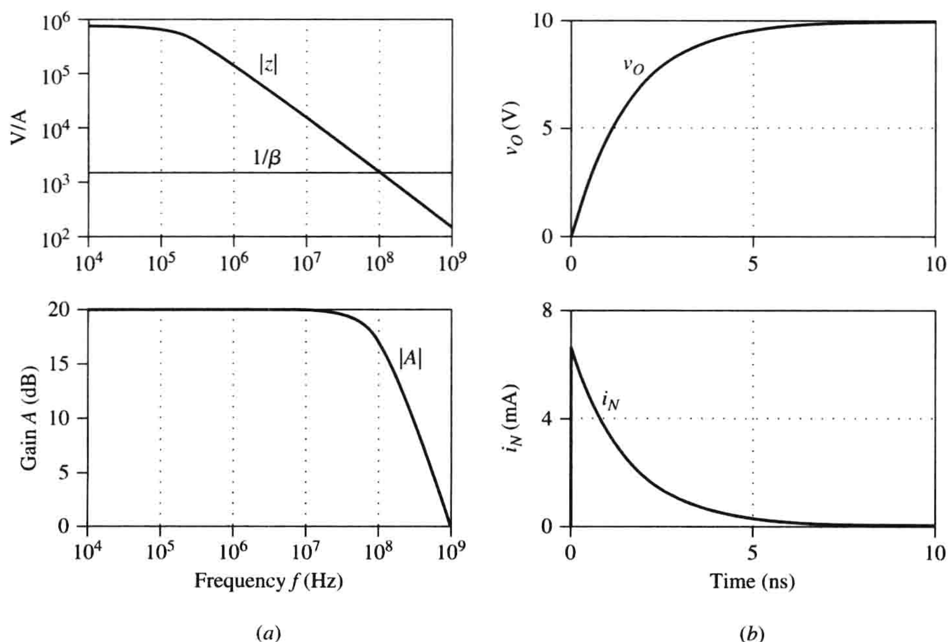


FIGURE 6.40
(a) Frequency response and (b) step response of the circuit of Fig. 6.39.

FIGURE 6.41
Simplified circuit diagram of a CFA-derived VFA.

voltage-feedback amplifiers (VFAs),[12] alongside the current-feedback amplifiers (CFAs) just discussed. Though the borderline between standard and high-speed VFAs keeps changing, at the time of writing we can take a high-speed VFA as one having[13] GBP > 50 MHz and SR > 100 V/μs. Two of the most popular high-speed VFA architectures in current use are illustrated in Figs. 6.41 and 6.42.

The VFA of Fig. 6.41 is similar to the CFA of Fig. 6.33, except for the addition of a unity-gain buffer (Q_{13} through Q_{16}) to raise the input impedance at node v_N, and the fact that the dynamics-controlling resistor R is now driven by the two input buffers. Since the current available to charge/discharge the gain-node capacitance C_{eq} is proportional to the magnitude of the input voltage difference as $(v_P - v_N)/R$, this VFA retains the slewing characteristics of a CFA. However, in all other respects

FIGURE 6.42
Simplified circuit diagram of a folded cascode bipolar VFA.

this architecture exhibits the characteristics of a VFA, namely, high input impedance at both nodes v_P and v_N, a decreasing closed-loop bandwidth with increasing closed-loop gain, and better dc characteristics than CFAs as the dc errors of the two matched input buffers tend to cancel each other out. This architecture can be used in all traditional VFA configurations, including inverting integrators. An example of a VFA using this architecture is the LT1363 70-MHz, 1000-V/μs op amp.

The trend toward high speed as well as low-power-supply voltages has inspired the folded cascode architecture, which finds wide use both in complementary bipolar processes and CMOS processes. In the bipolar illustration[14] of Fig. 6.42, any imbalance between v_P and v_N will cause an imbalance in the collector currents of the common-emitter *npn* pair Q_1 and Q_2, and this current imbalance is in turn fed to the emitters of the common-base *pnp* pair Q_3 and Q_4 (hence the term *folded cascode*). The latter pair is actively loaded by the current mirror Q_5 and Q_6 to provide high voltage gain at the gain node, whence the signal is buffered to the outside via a suitable unity-gain stage. Product examples utilizing this architecture are the EL2044C low-power/low-voltage 120-MHz unity-gain stable op amp, and the THS4401 high-speed VFA offering a unity-gain bandwidth of 300 MHz, SR = 400 V/μs, and t_S = 30 ns to 0.1%.

PROBLEMS

6.1 Open-loop frequency response

6.1 Cursor measurements on the 741 response of Fig. 6.1*b* give $a_0 = 185,200$ V/V, $f_b =$ 5.26 Hz, f_t = 870.1 kHz, and ph $a(jf_t) = -116.7°$. (*a*) Assuming all higher-order roots can be modeled with a single-pole frequency $f_p (f_p > f_t)$, estimate f_p so as to match the value of ph $a(j870.1 \text{ kHz})$. (*b*) Find the new values of f_t and ph $a(jf_t)$, and comment.

6.2 Given that the gain of a constant-GBP op amp has a magnitude of 80 dB at $f = 10$ Hz and a phase angle of $-58°$ at $f = 320$ Hz, estimate a_0, f_b, and f_t.

6.3 (*a*) Because of manufacturing process variations, the second-stage gain of a certain 741 op amp version is $-a_2 = -544$ V/V $\pm 20\%$. How does this affect a_0, f_b, and f_t? (*b*) Repeat, but for $C_c = 30$ pF $\pm 10\%$.

6.4 Given that a constant-GBP op amp has $|a(j100 \text{ Hz})| = 1$ V/mV and $|a(j1 \text{ MHz})| =$ 10 V/V, find (*a*) the frequency at which $\measuredangle a = -60°$, and (*b*) the frequency at which $|a| = 2$ V/V. *Hint:* Start out with the linearized magnitude plot.

6.2 Closed-loop frequency response

6.5 Show that the circuit of Example 6.2 yields $A(jf) = H_{0\text{LP}} \times H_{\text{LP}}$. What are the values of $H_{0\text{LP}}$, f_0, and Q?

6.6 (*a*) Show that cascading n identical noninverting amplifiers with individual dc gains A_0 yields a composite amplifier with the overall bandwidth $f_B = (f_t/A_0)\sqrt{2^{1/n} - 1}$. (*b*) Develop a similar expression for the case of n inverting amplifiers with individual dc gains $-A_0$.

6.7 (*a*) Repeat Example 6.2, but for a cascade of three 741 noninverting amplifiers with individual dc gains of 10 V/V. (*b*) Compare the −3-dB bandwidths of the one-op amp, two-op amp, and three-op amp designs, and comment.

6.8 (*a*) Consider the cascade connection of a noninverting amplifier with $A_0 = 2$ V/V, and an inverting amplifier with $A_0 = -2$ V/V. If both amplifiers use op amps with GBP = 5 MHz, find the −3-dB frequency of the composite amplifier. (*b*) Find the 1% magnitude error and the 5° phase-error bandwidths.

6.9 (*a*) Find the closed-loop GBP of the inverting amplifier of Fig. P1.64 if $R_1 = R_2 = \cdots = R_6 = R$, $r_d \gg R$, $r_o \ll R$, and $f_t = 4$ MHz. (*b*) Repeat if the source v_I is applied at the noninverting input and the left terminal of R_1 is connected to ground. (*c*) Repeat part (*b*), but with the left terminal of R_1 left floating. Comment.

6.10 (*a*) Using a 10-MHz op amp, design a two-input summing amplifier such that $v_O = -10(v_1 + v_2)$; hence, find its −3-dB frequency. (*b*) Repeat, but for five inputs, or $v_O = -10(v_1 + \cdots + v_5)$. Compare with the amplifier of part (*a*) and comment.

6.11 Assuming 741 op amps, find the −3-dB frequency of the circuits of (*a*) Fig. P1.17, (*b*) Fig. P1.19, (*c*) Fig. P1.22, and (*d*) Fig. P1.72.

6.12 Find the −3-dB frequency of the triple-op-amp IA of Fig. 2.21, given that all op amps have GBP = 8 MHz. Calculate with the wiper all the way down and all the way up.

6.13 In the dual-op amp IA of Fig. 2.23 let $R_3 = R_1 = 1$ kΩ, $R_4 = R_2 = 9$ kΩ, and $f_{t1} = f_{t2} = 1$ MHz. Find the −3-dB frequency with which the IA processes V_2, and that with which it processes V_1.

6.14 Sketch and label the frequency plot of the CMRR_{dB} of the IA of Problem 6.13. Except for the finite f_t, the op amps are ideal and the resistance ratios are perfectly matched.

6.15 A triple-op amp instrumentation amplifier with $A = 10$ V/V is to be designed using three constant-GBP, JFET-input op amps of the same family. Letting $A = A_{\text{I}} \times A_{\text{II}}$, how would you choose A_{I} and A_{II} in order to minimize the worst-case output dc error E_0? Maximize the overall −3-dB bandwidth?

6.16 Three signals v_1, v_2, and v_3 are to be summed using the topology of Fig. P1.33, and two alternatives are being considered: $v_O = v_1 + v_2 + v_3$ and $v_O = -(v_1 + v_2 + v_3)$. Which option is most desirable from the viewpoint of minimizing the untrimmed dc output error E_O? Maximizing the −3-dB frequency?

6.17 A unity-gain buffer is needed and the following options are being considered, each offering advantages and disadvantages in the event that the circuit must subsequently be altered: (*a*) a voltage follower, (*b*) a noninverting amplifier with $A_0 = 2$ V/V followed by a 2:1 voltage divider, and (*c*) a cascade of two unity-gain inverting amplifiers. Assuming constant-GBP op amps, compare the advantages and disadvantages of the three alternatives.

6.18 Find the closed-loop GBP of the inverting amplifier of Fig. 1.42*a*, given that $R_1 = 10\,\text{k}\Omega$, $R_2 = 20\,\text{k}\Omega$, $R_3 = 120\,\text{k}\Omega$, $R_4 = 30\,\text{k}\Omega$, and $f_t = 27$ MHz. Except for its finite f_t, the op amp can be considered ideal.

6.19 Find the closed-loop gain and bandwidth of the high-sensitivity I-V converter of Fig. 2.2 if $R = 200\ \text{k}\Omega$, $R_1 = R_2 = 100\ \text{k}\Omega$, and the input source has a 200-kΩ parallel resistance toward ground. The op amp is ideal, except for a constant GBP such that at 1.8 kHz the open-loop gain is 80 dB.

6.20 The circuit of Fig. P1.22 is implemented with three 10-kΩ resistances and an op amp with $a_0 = 50$ V/mV, $I_B = 50$ nA, $I_{OS} = 10$ nA, $V_{OS} = 0.75$ mV, $\text{CMRR}_{\text{dB}} = 100$ dB, and $f_t = 1$ MHz. Assuming $v_I = 5$ V, find the maximum dc output error as well as the small-signal bandwidth with both the switch open and the switch closed.

6.3 Input and output impedances

6.21 If the floating-load V-I converter of Fig. 2.4*a* is implemented with an op amp having $a_0 = 10^5$ V/V, $f_b = 10$ Hz, $r_d \gg R$, $r_o \ll R$, and $R = 10\ \text{k}\Omega$, sketch and label the magnitude Bode plot of the impedance $Z_o(jf)$ seen by the load; hence, find the element values of its equivalent circuit.

6.22 Find the impedance $Z_o(jf)$ seen by the load in the V-I converter of Fig. P2.7 if the op amp has $a_0 = 10^5$ V/V, $f_t = 1$ MHz, $r_d = \infty$, $r_o = 0$, $R_1 = R_2 = 18\ \text{k}\Omega$, and $R_3 = 2\ \text{k}\Omega$.

6.23 If the Howland current pump of Fig. 2.6*a* is implemented with four 10-kΩ resistances and an op amp having $a_0 = 10^5$ V/V, $f_t = 1$ MHz, $r_d = \infty$, and $r_o = 0$, sketch and label the magnitude plot of the impedance Z_o seen by the load. Justify using physical insight.

6.24 The negative-resistance converter of Fig. 1.21*b* is implemented with three 10-kΩ resistances and an op amp with GBP = 1 MHz. Find its impedance Z_{eq}. How does it change as f is swept from 0 to ∞?

6.25 The grounded-load current amplifier of Fig. 2.12*b* is implemented with $R_1 = R_2 = 10\ \text{k}\Omega$ and an op amp having $f_t = 10$ MHz, $r_d = \infty$, and $r_o = 0$. If the amplifier is driven by a source with a parallel resistance of 30 kΩ and drives a load of 2 kΩ, sketch and label the magnitude plots of the gain, the impedance seen by the source, and the impedance seen by the load.

6.26 A constant-GBP JFET-input op amp with $a_0 = 10^5$ V/V, $f_t = 4$ MHz, and $r_o = 100\ \Omega$ is configured as an inverting amplifier with $R_1 = 10\ \text{k}\Omega$ and $R_2 = 20\ \text{k}\Omega$. What is the frequency at which resonance with a 0.1-μF load capacitance will occur? What is the value of Q?

6.27 In the circuit of Fig. 1.42*a* let $R_1 = R_2 = R_3 = 30\ \text{k}\Omega$, $R_4 = \infty$, and let the op amp have $a_0 = 300$ V/mV and $f_b = 10$ Hz. Assuming $r_d = \infty$ and $r_o = 0$, sketch and label the magnitude plot of the impedance $Z(jf)$ between the node common to R_2 and R_3, and ground; use log-log scales.

6.28 In the circuit of Fig. 1.14*b* let both the 10-kΩ and 30-kΩ resistances be changed to 1 kΩ, and let the 20-kΩ resistance be changed to 18 kΩ. Assuming $r_d = \infty$, $r_o = 0$, and $f_t = 1$ MHz, sketch and label the magnitude plot of the impedance $Z(jf)$ seen by the input source; use log-log scales.

6.29 The op amp of Fig. 6.3*b* is configured as a unity-gain inverting amplifier via two identical 10-kΩ resistances. (*a*) Find the asymptotic values and the break frequencies

of the magnitude Bode plot of its output impedance $Z_o(jf)$. What are the element values of its equivalent circuit? (*b*) Repeat, but for the input impedance $Z_i(jf)$.

6.4 Transient response

6.30 Investigate the response of the high-sensitivity *I-V* converter of Example 2.2 to an input step of 10 nA. Except for $f_t = 1$ MHz and SR $= 5$ V/μs, the op amp is ideal.

6.31 Investigate the response of a Howland current pump to an input step of 1 V. The circuit is implemented with four 10-kΩ resistances and a 1-MHz op amp, and it drives a 2-kΩ load.

6.32 (*a*) Using a 741C op amp powered from ± 15-V regulated supplies, design a circuit that gives $v_O = -(v_I + 5\text{ V})$ with the maximum small-signal bandwidth possible. (*b*) What is this bandwidth? What is the FPB?

6.33 An inverting amplifier with $A_0 = -2$ V/V is driven with a square wave of peak values $\pm V_{im}$ and frequency f. With $V_{im} = 2.5$ V, it is observed that the output turns from trapezoidal to triangular when f is raised to 250 kHz; with $f = 100$ kHz, it is found that slew-rate limiting ceases when V_{im} is lowered to 0.4 V. If the input is changed to a 3.5-V (rms) ac signal, what is the useful bandwidth of the circuit? Is it small-signal or large-signal limited?

6.34 Find the response of the cascaded amplifier of Example 6.2 to a 1-mV input step.

6.35 A cascaded amplifier consists of an op amp OA_1, operating as a noninverting amplifier with $A_0 = +20$ V/V, followed by an op amp OA_2, operating as an inverting amplifier with $A_0 = -10$ V/V. Sketch the circuit; then find the minimum values of f_{t1}, SR_1, f_{t2}, and SR_2 needed to ensure an overall bandwidth of 100 kHz with a full-power output signal of 5 V (rms).

6.36 In the dual-op-amp IA of Fig. 2.23 let $R_3 = R_1 = 1\text{ k}\Omega$, $R_4 = R_2 = 9\text{ k}\Omega$, and $f_{t1} = f_{t2} =$ 1 MHz. Find the small-signal step response if (*a*) $v_1 = 0$ and the step is applied at v_2, (*b*) $v_2 = 0$ and the step is applied at v_1, and (*c*) the step is applied at v_1 and v_2 tied together.

6.37 (*a*) Using the LF353 dual JFET-input op amp, whose ratings are $V_{OS(\text{max})} = 10$ mV, GBP $= 4$ MHz, and SR $= 13$ V/μs, design a cascaded amplifier having an overall gain of 100 V/V as well as provision for overall offset-error nulling. (*b*) Find the small-signal bandwidth as well as the FPB. (*c*) If the circuit is to operate with a 50-mV (rms) ac input, what is its useful frequency range of operation? Is it small-signal or large-signal limited?

6.38 A JFET-input op amp is configured as an inverting amplifier with $A_0 = -10$ V/V and is driven by a 1-V (peak-to-peak) ac signal. Assuming $a_0 = 200$ V/mV, $f_t = 3$ MHz, and SR $= 13$ V/μs, estimate the peak-to-peak amplitude of the inverting input voltage v_N for $f = 1$ Hz, 10 Hz, ..., 10 MHz. Comment.

6.39 In the high-sensitivity *I-V* converter of Fig. 2.2 let $R = 100\text{ k}\Omega$, $R_1 = 10\text{ k}\Omega$, $R_2 = 30\text{ k}\Omega$, and let the op amp have $f_t = 4$ MHz and SR $= 15$ V/μs. Except for these limitations, the op amp can be considered ideal. If $i_I = 20\sin(2\pi ft)\ \mu$A, what is the useful bandwidth of the circuit? Is it small-signal or large-signal limited?

6.40 Equation (6.30) indicates that if we want to avoid slew-rate limiting in a voltage follower implemented with an op amp having SR = 0.5 V/μs and f_t = 1 MHz, we must limit the input step magnitude below about 80 mV. What is the maximum allowed input step if the same op amp is configured as: (*a*) An inverting amplifier with a gain of −1 V/V? (*b*) A noninverting amplifier with a gain of +2 V/V? (*c*) An inverting amplifier with a gain of −2 V/V?

6.41 Assuming equal resistors in the circuit of Fig. P1.64, find the minimum values of SR and f_t required for a useful bandwidth of 1 MHz for a sinusoidal input with a peak amplitude of 1 V.

6.42 The wideband band-pass filter of Example 3.5 is to be implemented with a constant GBP op amp. Find the minimum f_t and SR for an undistorted full-power output with a magnitude error of less than 1% over the entire audio range (that is, 20 Hz to 20 kHz).

6.5 Effect of finite GBP on integrator circuits

6.43 (*a*) Apply Eq. (6.13) to the integrator of Fig. 6.20*a* to calculate the zero pair of $H(s)$. Where are the two zeros located in the s-plane? *Hint*: In the frequency region of the zeros you can approximate $a(s) \cong \omega_t/s$ and $T(s) = a(s)\beta_\infty \cong a(s)$. (*b*) Find a transfer function $H_m(jf)$ that will match as closely as possible the response $H(jf)$ of the integrator of Fig. 6.20*a*. Start out with initial pole-frequency estimates of 100 Hz and 1 MHz, and zero-frequency estimates of 10 MHz each, and fiddle with their values until the PSpice plot of $|H_m(\text{jf})|$ matches that of $|H(\text{jf})|$ (you may wish to superimpose the respective plots to show how well they match).

6.44 Predict the transient response of the integrator of Fig. 6.20*a* to a 10-mV input step and compare with the integrator response in the limit $a \to \infty$.

6.45 Discuss the implications of replacing the op amp of Fig. 6.20*a* with a better unit having (*a*) f_t = 100 MHz, and (*b*) $a_0 = 10^4$ V/V. Sketch and label the magnitude plots of a, $1/\beta$, and H, and estimate the integrator's unity-gain frequency. Compare with Fig. 6.20*b* and comment.

6.46 The integrator of Fig. 3.6 is implemented with a JFET-input op amp having a_0 = 50 V/mV, f_t = 4 MHz, and r_o = 100Ω. If R = 15.8 kΩ and C = 1 nF, sketch and label the magnitude Bode plot of its output impedance $Z_o(jf)$. *Hint:* Use Eq. (6.23*b*).

6.47 A Deboo integrator is implemented with four 10-kΩ resistances, a 3.183-nF capacitance, and a 1-MHz op amp having a_0 = 106 dB, $r_d = \infty$, and $r_o = 0$. (*a*) Sketch and label the magnitude Bode plots of $a(jf)$, $1/\beta(jf)$, and $H(jf)$. (*b*) Estimate the downshift in the unity-gain frequency.

6.48 A Deboo integrator is implemented with four 10-kΩ resistances, a 1-nF capacitance, and a 1-MHz op amp having a_0 = 1 V/mV, $r_d = \infty$, and $r_o = 0$. Predict its transient response to a 10-mV input step and compare with the integrator response in the limit $a \to \infty$.

6.49 (*a*) Assuming $r_d = \infty$, $r_o = 0$, and $a(jf) \cong f_t/jf$, find $H(jf)$ for the compensated integrator of Fig. 6.21*a*. (*b*) Show that letting $C_c = C/(f_t/f_0 - 1)$ makes $H \cong H_{\text{ideal}}$. (*c*) Specify suitable components for $f_0 = 10$ kHz, and verify with PSpice for f_t = 1 MHz.

6.50 (*a*) Assuming $r_d = \infty$, $r_o = 0$, and $a(jf) \cong f_t/jf$, find $H(jf)$ for the compensated integrator of Fig. 6.21*b*. (*b*) Show that letting $R_c = 1/(2\pi C f_t)$ makes $H = H_{\text{ideal}}$. (*c*) Specify suitable components for $f_0 = 10$ kHz if $r_o = 100\ \Omega$, and verify with PSpice for $f_t = 1$ MHz.

6.51 (*a*) Find $H(jf)$ for the circuit of Fig. 6.22*b*, rationalize it, and discard higher-order terms to show that $\epsilon_\phi = +f/f_t$ for $f \ll f_t$. (*b*) Verify with PSpice for the case $f_0 = 10$ kHz and $f_t = 1$ MHz.

6.52 (*a*) Find an expression for the phase error of the Deboo integrator of Problem 6.47. (*b*) Find a suitable resistance R_c that, when placed in series with the capacitance, will provide phase-error compensation.

6.53 The active compensation scheme of Fig. P6.53 (see *Electronics and Wireless World*, May 1987) is a generalization of that of Fig. 6.22*a*, in that it allows for phase-error control. Verify that the error function of this circuit is $(1+jf/\beta_2 f_{t2})/(1+jf/f_{t1} - f^2/\beta_2 f_{t1} f_{t2})$, $\beta_2 = R_1/(R_1 + R_2)$. What happens if the op amps are matched and $R_1 = R_2$? Would you have any use for this circuit?

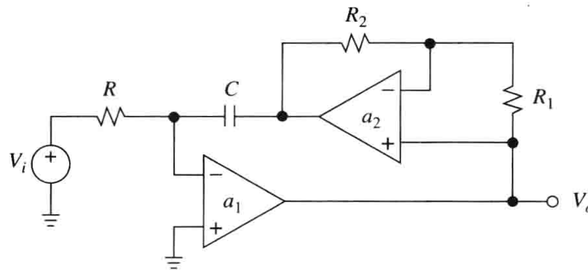


FIGURE P6.53

6.54 The active compensation method of Problem 6.53 can also be applied to the Deboo integrator, as shown in Fig. P6.54 (see *Proceedings of the IEEE*, February 1979, pp. 324–325). Show that for matched op amps and $f \ll f_t$ we have $\epsilon_\phi \cong -(f/0.5f_t)^3$.

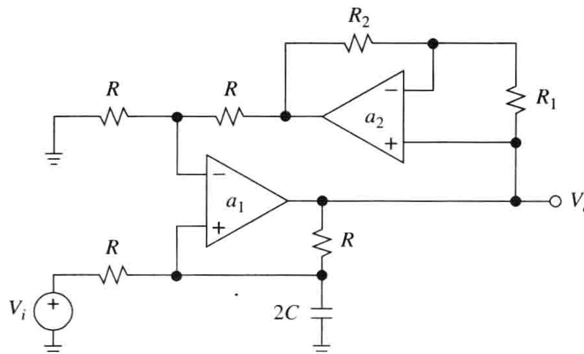


FIGURE P6.54

6.6 Effect of finite GBP on filters

6.55 Obtain an expression for the feedback factor $\beta(jf)$ of the low-pass circuit of Fig. 6.28*a*. What are the values of its pole and zero frequencies?

6.56 Sketch and label the magnitude plot of the output impedance $Z_o(jf)$ of the low-pass filter of Fig. 6.28*a*.

6.57 Obtain an expression for the feedback factor $\beta(jf)$ of the high-pass circuit of Fig. 6.29*a*. What are the values of its pole and zero frequencies?

6.58 Because of the op amp's finite GBP, the response of the high-pass filter of Fig. 6.29*a* is actually a band-pass, or $H(jf) = H_{0\text{BP}}H_{\text{BP}}(jf)$. (*a*) Assuming $r_d = \infty$ and $r_o = 0$, expand Eq. (6.50) to find expressions for H_{0BP}, Q, and the resonance frequency. (*b*) Calculate the lower -3-dB frequency of the band-pass response, compare it with the intended -3-dB frequency of the high-pass response, and comment.

6.59 Calculate the new values of $H_{0\text{BP}}$, Q, and f_0 if the circuit of Example 6.11 is implemented with an op amp having a low dc gain of $a_0 = 10$/V/V but extremely high bandwidth ($f_t \to \infty$). Comment on how the various parameters are affected.

6.60 Investigate the effect of a constant GBP op amp on the phase shifter of Fig. 3.12. What are the magnitude and phase errors at f_0 if $f_0 = 10$ kHz and GBP $= 1$ MHz?

6.61 (*a*) Specify suitable component values for the *D* element of Fig. 4.15*b* for a FDNR of –1 kΩ at 1 kHz. (*b*) Use PSpice to plot the FDNR versus frequency, both for the case of ideal op amps, and for the case of op amps with GBP = 1 MHz and dc open-loop gain of 100 dB (use log scales). What is the useful frequency range of the FDNR in the latter case?

6.62 Obtain an expression of the type of Eq. (6.51) for the low-pass *KRC* filter of Fig. 3.23.

6.63 Use the μA741 macromodel of PSpice to assess the departure from ideality of the band-pass response of the state-variable filter of Example 3.18. If needed, compensate and predistort to improve accuracy.

6.64 Investigate the effect of using an op amp with GBP $= 1$ MHz in the notch filter of Example 3.14.

6.65 The effect of finite GBP on the unity-gain *KRC* filter of Fig. 3.25 can be compensated for by placing a suitable resistance R_c in series with C and decreasing R to $R - R_c$. (*a*) Show that compensation is achieved for $R_c = 1/(2\pi C f_t)$. (*b*) Show the compensated circuit of Example 3.10 if the op amp is a 741 type.

6.7 Current-feedback amplifiers

6.66 In this and the following problems, assume a CFA with $z_0 = 0.5$ V/μA, $C_{\text{eq}} = 1.59$ pF, $r_n = 25$ Ω, $I_P = 1$ μA, $I_N = 2$ μA, and $(1/\beta)_{\min} = 1$ V/mA. Moreover, assume the input buffer has an offset voltage $V_{OS} = 1$ mV. (*a*) Using this CFA, design an inverting amplifier with $A_0 = -2$ V/V and the maximum possible bandwidth. What is this bandwidth? The dc loop gain? (*b*) Repeat, but for $A_0 = -10$ V/V and the same bandwidth as in part (*a*). (*c*) Repeat (*a*), but for a difference amplifier with a dc gain of 1 V/V.

6.67 (*a*) Using the CFA of Problem 6.66, design a voltage follower with the widest possible bandwidth. (*b*) Repeat, but for a unity-gain inverting amplifier. How do the closed-loop GBPs compare? (*c*) Modify both circuits so that the closed-loop bandwidth is reduced in half. (*d*) How do the maximum dc output errors compare in the various circuits?

6.68 (*a*) Using the CFA of Problem 6.66, provide two designs for an *I-V* converter with a dc sensitivity of −10 V/mA. (*b*) How do the closed-loop bandwidths compare? How do the maximum output errors compare?

6.69 The data sheets recommend the circuit of Fig. P6.69 to adjust the closed-loop dynamics. Assuming the CFA data of Problem 6.66, estimate the closed-loop bandwidth and rise time as the wiper is varied from end to end.

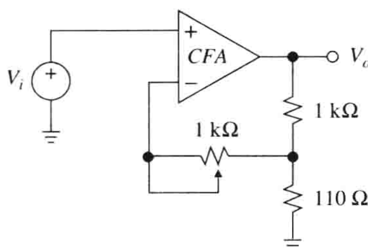


FIGURE P6.69

6.70 Using the CFA of Problem 6.66, design a second-order 10-MHz low-pass filter with $Q = 5$.

6.71 (*a*) Sketch a block diagram of the type of Fig. 6.34, but for the CFA-derived VFA of Fig. 6.41. Hence, denoting the output resistance of each input buffer as r_o, obtain expressions for the open-loop gain $a(jf)$ and the slew-rate SR. (*b*) Assuming $z(jf)$ can be modeled with a 1-MΩ resistance in parallel with a 2-pF capacitance, and $R = 500\ \Omega$ and $r_o = 25\ \Omega$, find a_0, f_b, f_t, β, T_0, A_0, and f_B, if $R_1 = R_2 = 1\ \text{k}\Omega$. (*c*) What is the SR for the case of a 1-V input step?

REFERENCES

1. J. E. Solomon, "The Monolithic Operational Amplifier: A Tutorial Study," *IEEE J. Solid-State Circuits,* Vol. SC-9, December 1974, pp. 314–332.
2. S. Franco, *Electric Circuits Fundamentals,* Oxford University Press, New York, 1995.
3. R. I. Demrow, "Settling Time of Operational Amplifiers," Application Note AN-359, *Applications Reference Manual,* Analog Devices, Norwood, MA, 1993.
4. C. T. Chuang, "Analysis of the Settling Behavior of an Operational Amplifier," *IEEE J. Solid-State Circuits,* Vol. SC-17, February 1982, pp. 74–80.
5. J. Williams, "Settling Time Measurements Demand Precise Test Circuitry," *EDN,* Nov. 15, 1984, p. 307.
6. P. R. Gray, P. J. Hurst, S. H. Lewis, and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 5th ed., John Wiley & Sons, New York, 2009, ISBN 978-0-470-24599-6.
7. P. O. Brackett and A. S. Sedra, "Active Compensation for High-Frequency Effects in Op Amp Circuits with Applications to Active RC Filters," *IEEE Trans. Circuits Syst.,* Vol. CAS-23, February 1976, pp. 68–72.

8. L. C. Thomas, "The Biquad: Part I—Some Practical Design Considerations," and "Part II—A Multipurpose Active Filtering System," *IEEE Trans. Circuit Theory,* Vol. CT-18, May 1971, pp. 350–361.
9. A. Budak, *Passive and Active Network Analysis and Synthesis,* Waveland Press, Prospect Heights, IL, 1991.
10. Based on the author's article "Current-Feedback Amplifiers Benefit High-Speed Designs," *EDN,* Jan. 5, 1989, pp. 161–172. © Cahners Publishing Company, a Division of Reed Elsevier Inc., Boston, 1997.
11. R. Mancini, "Converting from Voltage-Feedback to Current-Feedback Amplifiers," *Electronic Design Special Analog Issue,* June 26, 1995, pp. 37–46.
12. D. Smith, M. Koen, and A. F. Witulski, "Evolution of High-Speed Operational Amplifier Architectures," *IEEE J. Solid-State Circuits,* Vol. SC-29, October 1994, pp. 1166–1179.
13. Texas Instruments Staff, *DSP/Analog Technologies,* 1998 Seminar Series, Texas Instruments, Dallas, TX, 1998.
14. W. Kester, "High Speed Operational Amplifiers," *High Speed Design Techniques,* Analog Devices, Norwood, MA, 1996.
15. W. Jung, *Op Amp Applications Handbook* (Analog Devices Series), Elsevier/Newnes, Burlington, MA, 2005, ISBN 0-7506-7844-5.

7
NOISE

噪声通常是指任何会污损或干扰所关心信号的不希望的扰动。由输入偏置级电流和输入失调电压引起的失调误差就是我们熟知的噪声例子（这里的噪声指直流噪声）。然而，还有很多其他形式的噪声，特别是交流噪声。除非采取适当的降噪手段，否则噪声会明显降低电路的性能，根据噪声源的不同，可以将交流噪声分为外部噪声（或干扰噪声）和内部噪声（或固有噪声）。

干扰噪声

这种类型的噪声是由电路和外界之间，甚至是电路自身的不同部分之间多余的相互作用产生的。这种相互作用可以是电的、磁的、电磁的，甚至是机电的（比如拾音器噪声和压电噪声）。电的相互作用和磁的相互作用，是通过相邻电路之间或同一电路的相邻部分之间的寄生电容和互感产生的。电磁干扰的出现是因为每根导线和引线都构成了一个潜在的天线。外部噪声也可能会在无意

间，通过接地总线和电源总线进入电路。

干扰噪声可以是周期的、间歇的，或者完全随机的。将来自于电力线路的频率和它的谐波、无线电台、机械开关电弧、电抗元件电压尖脉冲等的静电与电磁噪声降到最小，可以降低或防止干扰噪声。这些防护措施包括滤波、去耦、隔离、静电和静电屏蔽、重新定位元件和引脚、采用消声器网络、消除接地回路和采用低噪声供电电源。虽然干扰噪声常常被误解成“不可捉摸的”，但它还是可以用一种合理的方法对它进行解释和处理的。

固有噪声

尽管能够设法消除全部干扰噪声，但是电路仍会呈现固有噪声。这种噪声形式本质上就是随机的。它源于各种随机现象，例如电阻中电子的热骚动，半导体中电子空穴对随机地产生和重组等。由于热骚动，电阻中每个振动的电子都会形成一个极小的电流。将这些电流进行代数累加，就形成了净电流和由此产生的电压。虽然净电压均值为零，但是由于单个电流瞬时幅度和方向都是随机分布的，所以净电压会不断波动。即使把电阻静静地放在抽屉里，这些波动仍然会发生。因此，我们可以假设电路中的每个节点电压和每个支路电流都是在它们的期望值附近不断地波动。

信噪比

噪声的存在会降低信号的质量，最终限制了能成功检测、测量和解释的信号大小。可以用信噪比（SNR）

$$\mathrm{SNR} = 10\log_{10}\frac{X_s^2}{X_n^2} \tag{7.1}$$

来表示在噪声存在的条件下，信号的质量。式中，X_s 是信号的均方根（rms）值，X_n 是噪声分量的 rms 值。SNR 越小，就越难从噪声中恢复有用信号。尽管经过适当的信号处理过程（比如信号平均），可以将湮没在噪声中的信号恢复出来，但是要让 SNR 像其他的设计限制条件一样尽量高。

电路设计者对噪声的关注程度最终依赖于应用对性能的要求。随着运算放大器输入失调误差性能的巨大改善，以及 A-D 和 D-A 转换器分辨率的大幅提高，噪声在高性能系统的误差预算分析中成为越来越重要的因素。下面举一个 12 位系统的例子，注意，当最大标定值为 10V 时，LSB /2（最低有效位的一半）对应于 $10/2^{12} = 1.22\text{mV}$，它自身可能就会给转换器设计带来问题。在实际环境中，由传感器产生的信号要求得到大幅的放大，以达到 10V 的最大标定值。设典型的满刻度传感器输出为 10mV，这里 LSB /2 等于 1.22μV。如果放大器仅产

生 1μV 的输入参考噪声，就会使 LSB 分辨率失效！

为了充分利用高档精密的设备和系统，设计者必须能够理解噪声的机理；对噪声进行计算，仿真和测量；根据要求使噪声最小。本章将会对这些问题进行介绍。

本章重点

本章先从介绍噪声的概念、计算、测量方法和频谱入手。然后探讨噪声的动态特性，其中，一些实用工具会重点介绍，包括分段图解积分、粉噪声的切线原理以及 PSpice 仿真。紧接着本章会讲述一些常见的噪声源，包括二极管、BJT、JFET 和 MOSFET 的噪声模型。最后，本章将会应用上述材料，探讨在电压反馈和电流反馈下放大器电路的噪声特性。在噪声不可以忽略的电路中，本章将会着重介绍光敏二极管放大器的噪声处理，包括图解噪声计算和噪声滤波。除此之外，本章还囊括了低噪声运算放大器。

噪声本身就是一个庞大的主题，已经有基本专门论述噪声的参考书[1,2]。出于必要性的原则，本书将讨论的范围限制在噪声的概念以及运算放大器用户关注的主要问题内。

7.1 NOISE PROPERTIES

Since noise is a random process, the instantaneous value of a noise variable is unpredictable. However, we can deal with noise on a statistical basis. This requires introducing special terminology as well as special calculation and measurement.

Rms Value and Crest Factor

Using subscript n to denote noise quantities, we define the *root-mean-square (rms) value* X_n of a noise voltage or current $x_n(t)$ as

$$X_n = \left(\frac{1}{T} \int_0^T x_n^2(t)\, dt \right)^{1/2} \tag{7.2}$$

where T is a suitable averaging time interval. The square of the rms value, or X_n^2, is called the *mean square value.* Physically, X_n^2 represents the average power dissipated by $x_n(t)$ in a 1-Ω resistor.

In voltage-comparator applications, such as A-D converters and precision multivibrators, accuracy and resolution are affected by the instantaneous rather than the rms value of noise. In these situations, expected peak values of noise are of more concern. Most noise has a Gaussian, or normal, distribution as shown in Fig. 7.1, so it is possible to predict instantaneous values in terms of probabilities. The *crest factor*

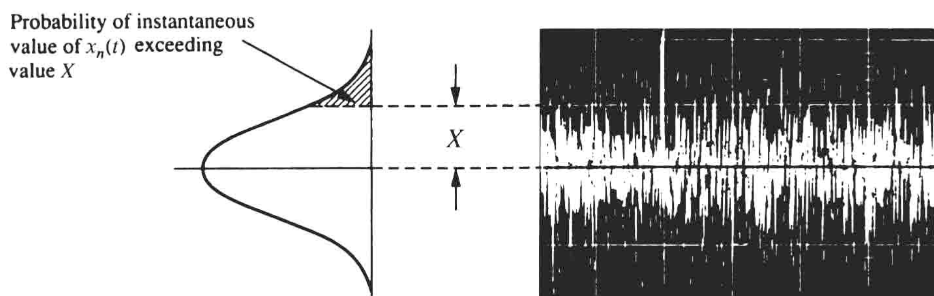


FIGURE 7.1
Voltage noise (right), and Gaussian distribution of amplitude.

(CF) is defined as the *ratio of the peak value to the rms value* of noise. Though all CF values are possible in principle, the likelihood of $x_n(t)$ exceeding a given value X decreases very rapidly with X, as indicated by the residual area under the distribution curve. Suitable calculations[6] reveal that for Gaussian noise the probability of CF exceeding 1 is 32%, that of exceeding 2 is 4.6%, that of exceeding 3 is 0.27%, that of exceeding 3.3 is 0.1%, and that of exceeding 4 is 0.0063%. It is common practice to take the *peak-to-peak value* of Gaussian noise to be 6.6 times the rms value, since the instantaneous value is within this range 99.9% of the time, which is close to 100%.

Noise Observation and Measurement

Voltage noise can readily be observed with an oscilloscope of adequate sensitivity. An advantage of this instrument is that it allows us to actually see the signal and thus make sure it is internal noise and not externally induced noise, such as 60-Hz pickup. One way of estimating the rms value is by observing the maximum peak-to-peak fluctuation, and then dividing by 6.6. A less subjective alternative[7] is to observe noise with two equally calibrated channels, and adjust the offset of one channel until the two noisy traces just merge; if we then remove both noise sources and measure the difference between the two clean traces, the result is approximately *twice the rms value*.

Noise can be measured with a multimeter. Ac meters fall into two categories: *true rms meters* and *averaging-type meters*. The former yield the correct rms value regardless of the waveform, provided that the CF specifications of the instrument are not exceeded. The latter are calibrated to give the rms value of a sine wave. They first rectify the signal and compute its average, which for ac signals is $2/\pi$ times the peak value; then they synthesize the rms value, which for ac signals is $1/\sqrt{2}$ times the peak value, by amplifying the average value by $(1/\sqrt{2})/(2/\pi) = 1.11$. For Gaussian noise the rms value is $\sqrt{\pi/2} = 1.25$ times the average value,[2] so the noise reading provided by an averaging-type meter must be multiplied by $1.25/1.11 = 1.13$, or, equivalently, it must be increased by $20\log_{10} 1.13 \cong 1$ dB to obtain the correct value.

Noise Summation

In noise analysis one often needs to find the rms value of noise voltages in series or noise currents in parallel. Given two noise sources $x_{n1}(t)$ and $x_{n2}(t)$, the mean square value of their sum is

$$X_n^2 = \frac{1}{T}\int_0^T [x_{n1}(t) + x_{n2}(t)]^2\, dt = X_{n1}^2 + X_{n2}^2 + \frac{2}{T}\int_0^T x_{n1}(t)x_{n2}(t)\, dt$$

If the two signals are uncorrelated, as is usually the case, the average of their product vanishes, so the rms values add up in Pythagorean fashion,

$$X_n = \sqrt{X_{n1}^2 + X_{n2}^2} \tag{7.3}$$

This indicates that if the sources are of uneven strengths, minimization efforts should be directed primarily at the strongest one. For instance, two noise sources with rms values of 10 μV and 5 μV combine to give an overall rms value of $\sqrt{10^2 + 5^2} =$ 11.2 μV, which is only 12% higher than that of the dominant source. It is readily seen that reducing the dominant source by 13.4% has the same effect as eliminating the weaker source altogether!

As mentioned, the dc error referred to the input is also a form of noise, so when performing budget-error analysis we must add dc noise and rms ac noise *quadratically*.

Noise Spectra

Since X_n^2 represents the average power dissipated by $x_n(t)$ in a 1-Ω resistor, the physical meaning of mean square value is the same as for ordinary ac signals. However, unlike an ac signal, whose power is concentrated at just one frequency, noise power is usually spread all over the frequency spectrum because of the random nature of noise. Thus, when referring to rms noise, we must always specify the *frequency band* over which we are making our observations, measurements, or calculations.

In general, noise power depends on both the width of the frequency band and the band's location within the frequency spectrum. The rate of change of noise power with frequency is called the *noise power density*, and is denoted as $e_n^2(f)$ in the case of voltage noise, and $i_n^2(f)$ in the case of current noise. We have

$$e_n^2(f) = \frac{dE_n^2}{df} \qquad i_n^2(f) = \frac{dI_n^2}{df} \tag{7.4}$$

where E_n^2 and I_n^2 are the mean square values of voltage noise and current noise. Note that the units of $e_n^2(f)$ and $i_n^2(f)$ are volts squared per hertz (V²/Hz) and amperes squared per hertz (A²/Hz). Physically, noise power density represents the average noise power over a 1-Hz bandwidth as a function of frequency. When plotted versus frequency, it provides a visual indication of how power is distributed over the frequency spectrum. In integrated circuits, the two most common forms of power density distribution are white noise and $1/f$ noise.

The quantities $e_n(f)$ and $i_n(f)$ are called the *spectral noise densities,* and are expressed in volts per square root of hertz (V/$\sqrt{\text{Hz}}$) and amperes per square

root of hertz (A/$\sqrt{\text{Hz}}$). Some manufacturers specify noise in terms of noise power densities, others in terms of spectral noise densities. Conversion between the two is accomplished by squaring or by extracting the square root.

Multiplying both sides in Eq. (7.4) by df and integrating from f_L to f_H, the lower and upper limits of the frequency band of interest, allows us to find the rms values in terms of the power densities,

$$E_n = \left(\int_{f_L}^{f_H} e_n^2(f)\,df\right)^{1/2} \qquad I_n = \left(\int_{f_L}^{f_H} i_n^2(f)\,df\right)^{1/2} \tag{7.5}$$

Once again it is stressed that the concept of rms cannot be separated from that of frequency band: in order to find the rms value, we need to know the lower and upper limits of the band as well as the density within the band.

White Noise and 1/*f* Noise

White noise is characterized by a uniform spectral density, or $e_n = e_{nw}$ and $i_n = i_{nw}$, where e_{nw} and i_{nw} are suitable constants. It is so called by analogy with white light, which consists of all visible frequencies in equal amounts. When played through a loudspeaker, it produces a waterfall sound. Applying Eq. (7.5) we get

$$E_n = e_{nw}\sqrt{f_H - f_L} \qquad I_n = i_{nw}\sqrt{f_H - f_L} \tag{7.6}$$

indicating that the rms value of white noise increases with the square root of the frequency band. For $f_H \geq 10f_L$ we can approximate as $E_n \cong e_{nw}\sqrt{f_H}$ and $I_n \cong i_{nw}\sqrt{f_H}$ at the risk of an error of about 5% or less.

Squaring both sides in Eq. (7.6) yields $E_n^2 = e_{nw}^2(f_H - f_L)$ and $I_n^2 = i_{nw}^2(f_H - f_L)$, indicating that white-noise power is *proportional to the bandwidth,* regardless of the band's location within the frequency spectrum. Thus, the noise power within the 10-Hz band between 20 Hz and 30 Hz is the same as that within the band between 990 Hz and 1 kHz.

The other common form of noise is $1/f$ noise, so called because its power density varies with frequency as $e_n^2(f) = K_v^2/f$ and $i_n^2(f) = K_i^2/f$, where K_v and K_i are suitable constants. The spectral densities are $e_n = K_v/\sqrt{f}$ and $i_n = K_i/\sqrt{f}$, indicating that when plotted versus frequency on logarithmic scales, power densities have a slope of -1 dec/dec, and spectral densities a slope of -0.5 dec/dec. Substituting into Eq. (7.5) and integrating yields

$$E_n = K_v\sqrt{\ln(f_H/f_L)} \qquad I_n = K_i\sqrt{\ln(f_H/f_L)} \tag{7.7}$$

Squaring both sides in Eq. (7.7) yields $E_n^2 = K_v^2\ln(f_H/f_L)$ and $I_n^2 = K_i^2\ln(f_H/f_L)$, indicating that $1/f$-noise power is *proportional to the log ratio of the frequency band extremes,* regardless of the band's location within the frequency spectrum. Consequently, $1/f$ noise is said to have the same power content in each frequency decade (or octave). Once the noise rms of a particular decade (or octave) is known, the noise rms over m decades (or octaves) is obtained by multiplying the former by $\sqrt{m}$. For example, if the rms value within the decade 1 Hz $\leq f \leq$ 10 Hz is 1 μV, then the noise rms in the 9-decade span below 1 Hz, that is, down to about 1 cycle per 32 years, is $\sqrt{9} \times 1\ \mu\text{V} = 3\ \mu\text{V}$.

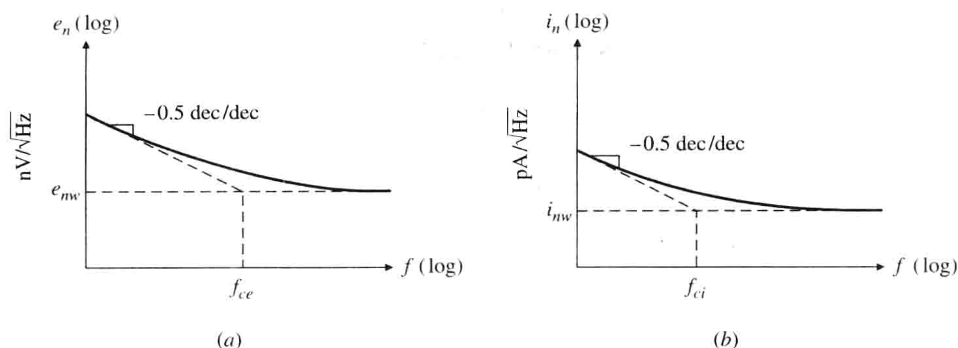


FIGURE 7.2
Typical IC noise densities.

Integrated-Circuit Noise

Integrated-circuit noise is a mixture of white and $1/f$ noise, as shown in Fig. 7.2. At high frequencies, noise is predominantly white, while at low frequencies $1/f$ noise dominates. The borderline frequency, or *corner frequency,* is found graphically as the intercept of the $1/f$ asymptote and the white-noise floor. Power densities are expressed analytically as

$$e_n^2 = e_{nw}^2\left(\frac{f_{ce}}{f}+1\right) \qquad i_n^2 = i_{nw}^2\left(\frac{f_{ci}}{f}+1\right) \tag{7.8}$$

where e_{nw} and i_{nw} are the *white-noise floors,* and f_{ce} and f_{ci} the *corner frequencies*. The μA741 data sheets of Fig. 5A.8 indicate $e_{nw} \cong 20$ nV/$\sqrt{\text{Hz}}$, $f_{ce} \cong 200$ Hz, $i_{nw} \cong 0.5$ pA/$\sqrt{\text{Hz}}$, and $f_{ci} \cong 2$ kHz. Inserting Eq. (7.8) into Eq. (7.5) and integrating, we get

$$E_n = e_{nw}\sqrt{f_{ce}\ln(f_H/f_L) + f_H - f_L} \tag{7.9a}$$

$$I_n = i_{nw}\sqrt{f_{ci}\ln(f_H/f_L) + f_H - f_L} \tag{7.9b}$$

EXAMPLE 7.1. Estimate the rms input voltage noise of the 741 op amp over the following frequency bands: (*a*) 0.1 Hz to 100 Hz (instrumentation range), (*b*) 20 Hz to 20 kHz (audio range), and (*c*) 0.1 Hz to 1 MHz (wideband range).

Solution.

(*a*) Equation (7.9*a*) gives $E_n = 20\times10^{-9}\sqrt{200\ln(10^2/0.1)+10^2-0.1} = 20\times10^{-9}\sqrt{1382+98.9} = 0.770\ \mu\text{V}$
(*b*) $E_n = 20\times10^{-9}\sqrt{1382+19{,}980} = 2.92\ \mu\text{V}$
(*c*) $E_n = 20\times10^{-9}\sqrt{3224+10^6} = 20.0\ \mu\text{V}$

We observe that $1/f$ noise dominates at low frequencies, white noise dominates at high frequencies—and the wider the frequency band, the higher the noise. Consequently, to minimize noise one should *limit the bandwidth to the strict minimum required.*

7.2 NOISE DYNAMICS

A common task in noise analysis is finding the total rms noise at the output of a circuit, given the noise density at its input as well as its frequency response. A typical example is offered by the voltage amplifier. The noise density at the output is $e_{no}(f) = |A_n(jf)|e_{ni}(f)$, where $e_{ni}(f)$ is the noise density at the input and $A_n(jf)$ is the *noise gain*. The *total output rms noise above* f_L is then $E_{no}^2 = \int_{f_L}^{\infty} e_{no}^2(f)\,df$, or

$$E_{no}^2 = \left(\int_{f_L}^{\infty} |A_n(jf)|e_{ni}^2(f)\,df\right)^{1/2} \tag{7.10a}$$

Another common example is offered by the transimpedance amplifier. Denoting its input noise density as $i_n(f)$ and its transimpedance noise gain as $Z_n(if)$, we have in this case

$$E_{no}^2 = \left(\int_{f_L}^{\infty} |Z_n(jf)|i_{ni}^2(f)\,df\right)^{1/2} \tag{7.10b}$$

(You can readily extend the above concepts to the other amplifier types, namely, current amplifiers and transadmittance amplifiers.) In the case of white noise, it is often convenient to let $f_L \to 0$ to simplify the math.

Noise Equivalent Bandwidth (NEB)

As an application example of Eq. (7.10), consider the case of white noise with spectral density e_{nw} going through a simple *RC* filter as in Fig. 7.3*a*. Since $|A_n|^2 = 1/[1 + (f/f_0)^2]$, where f_0 is the −3-dB frequency, Eq. (7.10*a*) gives (see Problem 7.3)

$$E_{no} = e_{nw}\left(\int_{f_L}^{\infty} \frac{df}{1+(f/f_0)^2}\right)^{1/2} = e_{nw}\sqrt{\text{NEB}}$$

where

$$\text{NEB} = f_0\frac{\pi}{2} - f_0 \tan^{-1}\frac{f_L}{f_0} \tag{7.11a}$$

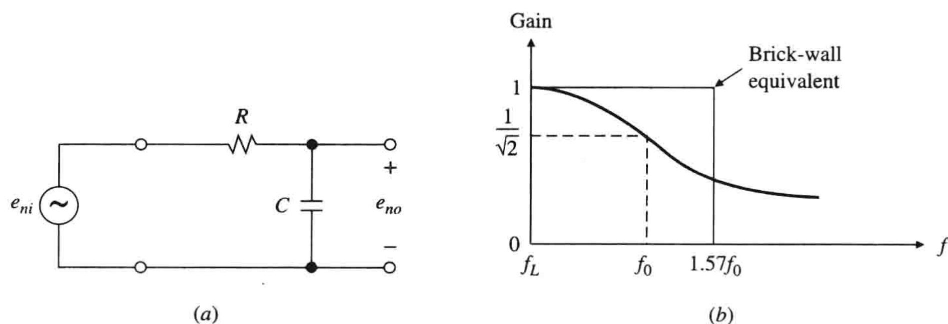


FIGURE 7.3
Noise equivalent bandwidth (NEB).

is called the *white-noise equivalent bandwidth*. We visualize it as the bandwidth of a brick-wall band-pass filter having a lower cutoff frequency of $f_{\text{lower}} = f_0 \tan^{-1}(f_L/f_0)$, and an upper cutoff frequency of $f_{\text{higher}} = f_0\pi/2 = 1.57 f_0$. Most cases of practical interest are such that $f_L \ll f_0$, so using the fact that $\tan^{-1} x \cong x$ for $x \ll 1$, we approximate $f_{\text{lower}} \cong f_0 f_L/f_0 = f_L$. Rewriting as

$$\text{NEB} \cong 1.57 f_0 - f_L \tag{7.11b}$$

and using Fig. 7.3*b* as reference, we observe that the fraction of 0.57 accounts for the transmitted noise above f_0 as a consequence of the gradual rolloff, or *skirt*. This property holds for all first-order low-pass functions, not just for *RC* networks. As we know, the closed-loop response of many amplifiers is a first-order function with $f_B = \beta f_t$ as the −3-dB frequency. These amplifiers pass white noise with a cutoff frequency of $1.57 f_B$.

More generally, the NEB of a circuit with noise gain $A_n(jf)$ is defined, for $f_L = 0$, as[2]

$$\text{NEB} = \frac{1}{A^2_{n(\max)}} \int_0^\infty |A_n(jf)|^2 \, df \tag{7.12}$$

where $A_{n(\max)}$ is the peak magnitude of the noise gain. The NEB represents the *frequency span of a brick-wall power gain response having the same area as the power gain response of the original circuit.*

The NEB can be computed analytically for higher-order responses. For instance, for an *n*th-order maximally flat low-pass response we have, for $f_L = 0$,

$$\text{NEB}_{\text{MF}} = \int_0^\infty \frac{df}{1 + (f/f_0)^{2n}} \tag{7.13a}$$

The results are[2] $\text{NEB}_{\text{MF}} = 1.57 f_0$ for $n = 1$, $1.11 f_0$ for $n = 2$, $1.05 f_0$ for $n = 3$, and $1.025 f_0$ for $n = 4$, indicating that NEB_{MF} rapidly approaches f_0 as n is increased.

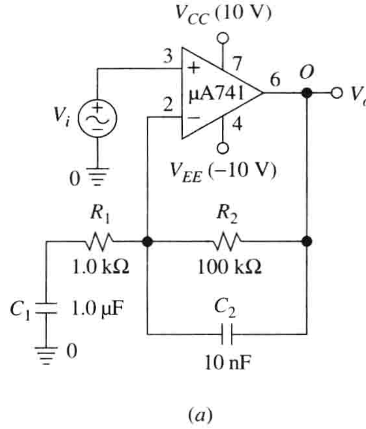
Likewise, it can be proven[5] that the noise equivalent bandwidths of the standard second-order low-pass and band-pass functions H_{LP} and H_{BP} defined in Section 3.4 are, respectively,

$$\text{NEB}_{\text{LP}} = Q^2 \text{NEB}_{\text{BP}} = Q\pi f_0/2 \tag{7.13b}$$

When the NEB cannot be calculated analytically, it can be estimated by piecewise graphical integration, or it can be found by computer via numerical integration.

EXAMPLE 7.2. Using PSpice, find the NEB of the 741 circuit of Fig. 7.4*a* for $f_L = 1$ Hz.

Solution. The gain plot of Fig. 7.4*b* (top) indicates that $|A|$ peaks at 51 V/V. Then, to find the NEB, we use PSpice's "s" function to calculate the squared area under the $|A|$ curve, and we divide the result by $51^2 = 2601$, in accordance with the definition of Eq. (7.12). The result is shown in Fig. 7.4*b* (bottom). Using PSpice's cursor facility, we measure $f_H = 1052$ Hz, so NEB $= 1052 - 1 = 1051$ Hz.



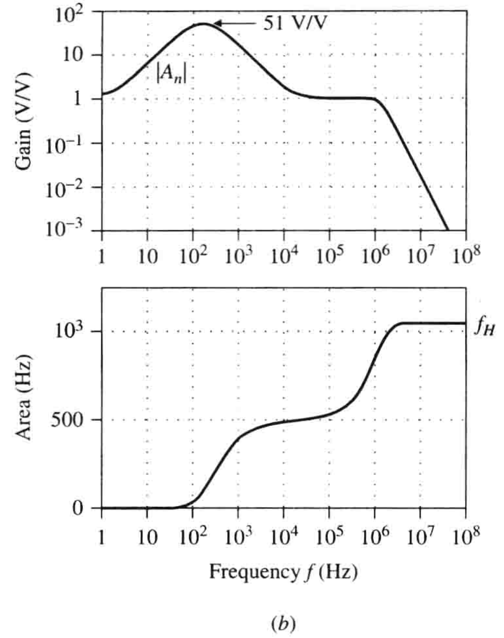


FIGURE 7.4
(a) PSpice circuit of Example 7.2 and (b) its voltage gain $|A|$ (top). To find the NEB, we direct PSpice to plot the function `(S(Vm(O)*Vm(O))/2601` (bottom).

Upper-Brick-Wall Frequency for 1/f Noise

By analogy with white noise, we seek an expression for f_H when we pass $1/f$ noise through a first-order low-pass filter having f_0 as its −3-dB frequency. Expressing power density as $e_{ni}^2(f) = K_v^2/f$, we adapt Eq. (7.10a) to find (see Problem 7.3) the *total output rms noise above* f_L as

$$E_{no} = K_v \left(\int_{f_L}^{\infty} \frac{df}{f[1+(f/f_0)^2]} \right)^{1/2} = K_v \sqrt{\ln \frac{f_H}{f_L}}$$

where

$$f_H = f_0 \sqrt{1 + (f_L/f_0)^2} \tag{7.14a}$$

is the upper-cutoff frequency of a brick-wall band-pass filter having a lower-cutoff frequency of f_L. Most cases of practical interest are such that $f_L << f_0$, so henceforth we shall simply let

$$f_H \cong f_0 \tag{7.14b}$$

Piecewise Graphical Integration

Noise densities and noise gains are often available only in graphical form. When this is the case, E_{no} is estimated by graphical integration, as illustrated in the following example.

EXAMPLE 7.3. Estimate the total rms output noise above 1 Hz for noise with the spectral density of Fig. 7.5 (top) going through an amplifier with the noise-gain characteristic of Fig. 7.5 (center).

Solution. To find the output density e_{no}, we multiply out the two curves point by point and obtain the curve of Fig. 7.5 (bottom). Clearly, the use of linearized Bode plots simplifies graphical multiplications considerably. Next, we integrate e_{no}^2 from $f_L = 1$ Hz to $f_H = \infty$. To facilitate our task, we break down the integration interval into three parts, as follows.

For 1 Hz $\leq f \leq$ 1 kHz we can apply Eq. (7.9*a*) with $e_{nw} = 20$ nV/$\sqrt{\text{Hz}}$, $f_{ce} =$ 100 Hz, $f_L = 1$ Hz, and $f_H = 1$ kHz. The result is $E_{no1} = 0.822\ \mu$V.

For 1 kHz $\leq f \leq$ 10 kHz the density e_{no} increases with f at the rate of +1 dec/dec, so we can write $e_{no}(f) = (20\ \text{nV}/\sqrt{\text{Hz}}) \times (f/10^3) = 2 \times 10^{-11} f$ V/$\sqrt{\text{Hz}}$. Then,

$$E_{no2} = 2 \times 10^{-11}\left(\int_{10^3}^{10^4} f^2\, df\right)^{1/2} = 2 \times 10^{-11}\left(\tfrac{1}{3} f^3\Big|_{10^3}^{10^4}\right)^{1/2} = 11.5\ \mu\text{V}$$

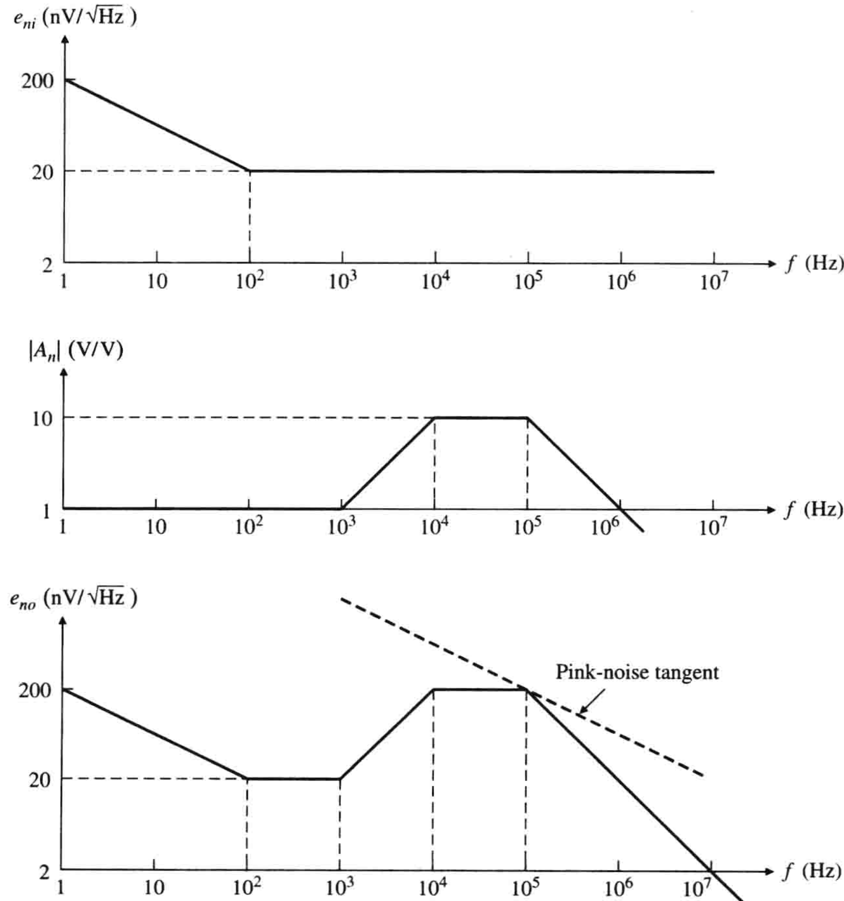


FIGURE 7.5
Noise spectra of Example 7.3.

For 10 kHz $\leq f \leq \infty$ we have white noise with $e_{nw} = 200\ \text{nV}/\sqrt{\text{Hz}}$ going through a low-pass filter with $f_0 = 100$ kHz. By Eq. (7.13*b*), $E_{no3} = 200 \times 10^{-9}(1.57 \times 10^5 - 10^4)^{1/2} = 76.7\ \mu\text{V}$.

Finally, we add up all components in rms fashion to obtain $E_{no} = \sqrt{E_{no1}^2 + E_{no2}^2 + E_{no3}^2} = \sqrt{0.822 + 11.5^2 + 76.7^2} = 77.5\ \mu\text{V}$.

The Pink-Noise Tangent Principle

Looking at the result of the foregoing example, we note that the largest contribution comes from E_{no3}, which represents noise above 10 kHz. We wonder if there is a quick method of predicting this, without having to go through all calculations. Such a method exists; it is offered by the *pink-noise tangent principle.*[6]

The pink-noise curve is the locus of points contributing equal-per-decade (or equal-per-octave) noise power. Its noise density slope is −0.5 dec/dec. The pink-noise principle states that if we lower the pink-noise curve until it becomes tangent to the noise curve $e_{no}(f)$, then the main contribution to E_{no} will come from the portions of the noise curve in the immediate vicinity of the tangent. In the example of Fig. 7.6 (bottom), the portions closest to the tangent are those leading to E_{no3}. We could have set $E_{no} \cong E_{no3} = 76.7\ \mu\text{V}$ without bothering to calculate E_{no1} and E_{no2}. The error caused by this approximation is insignificant, especially in light of the spread in noise data due to production variations. As we proceed, we shall make frequent use of this principle.

7.3 SOURCES OF NOISE

For an effective selection and utilization of integrated circuits, the system designer needs to be familiar with the basic noise-generating mechanisms in semiconductor devices. A brief discussion of these mechanisms follows.

Thermal Noise

Thermal noise, also called *Johnson noise,* is present in all passive resistive elements, including the stray series resistances of practical inductors and capacitors. Thermal noise is due to the random thermal motion of electrons (or holes, in the case of *p*-type semiconductor resistors). It is unaffected by dc current, so a resistor generates thermal noise even when sitting in a drawer.

As shown in Fig. 7.6*a*, thermal noise is modeled by a noise voltage of spectral density e_R in *series* with an otherwise noiseless resistor. Its power density is

$$e_R^2 = 4kTR \tag{7.15a}$$

where $k = 1.38 \times 10^{-23}$ J/K is Boltzmann's constant, and T is absolute temperature, in kelvins. At 25 °C, $4kT = 1.65 \times 10^{-20}$ W/Hz. An easy figure to remember is that at 25 °C, $e_R \cong 4\sqrt{R}\ \text{nV}/\sqrt{\text{Hz}}$, R in kilo-ohms. For instance, $e_{100\ \Omega} = 4\sqrt{0.1} = 1.26\ \text{nV}/\sqrt{\text{Hz}}$, and $e_{10\ \text{k}\Omega} = 12.6\ \text{nV}/\sqrt{\text{Hz}}$.

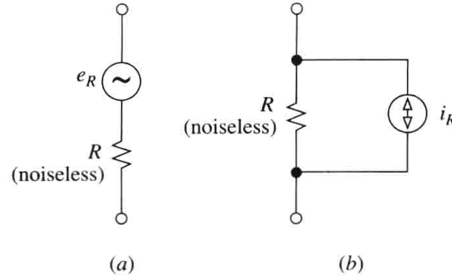


FIGURE 7.6
Thermal noise models.

Converting from Thévenin to Norton, we can model thermal noise also with a noise current i_R in *parallel* with an otherwise noiseless resistor, as shown in Fig. 7.6*b*. We have $i_R^2 = e_R^2/R^2$, or

$$i_R^2 = 4kT/R \tag{7.15b}$$

The preceding equations indicate that thermal noise is of the white type. Purely reactive elements are free from thermal noise.

EXAMPLE 7.4. Consider a 10-kΩ resistor at room temperature. Find (*a*) its voltage and (*b*) current spectral densities, and (*c*) its rms noise voltage over the audio range.

Solution.

(*a*) $e_R = \sqrt{4kTR} = \sqrt{1.65 \times 10^{-20} \times 10^4} = 12.8\ \text{nV}/\sqrt{\text{Hz}}$
(*b*) $i_R = e_R/R = 1.28\ \text{pA}/\sqrt{\text{Hz}}$
(*c*) $E_R = e_R\sqrt{f_H - f_L} = 12.8 \times 10^{-9} \times \sqrt{20 \times 10^3 - 20} = 1.81\ \mu\text{V}$

Shot Noise

This type of noise arises whenever charges cross a potential barrier, such as in diodes or transistors. Barrier crossing is a purely random event and the dc current we observe microscopically is actually the sum of many random elementary current pulses. Shot noise has a uniform power density,

$$i_n^2 = 2qI \tag{7.16}$$

where $q = 1.602 \times 10^{-19}$ C is the electron charge, and I is the dc current through the barrier. Shot noise is present in BJT base currents as well as in current-output D-A converters.

EXAMPLE 7.5. Find the signal-to-noise ratio for diode current over a 1-MHz bandwidth if (*a*) $I_D = 1\ \mu\text{A}$ and (*b*) $I_D = 1$ nA.

Solution.

(*a*) $I_n = \sqrt{2qI_Df_H} = \sqrt{2 \times 1.62 \times 10^{-19} \times 10^{-6} \times 10^6} = 0.57$ nA (rms). Thus, SNR = $20\log_{10}[(1\ \mu\text{A})/(0.57\ \text{nA})] = 64.9$ dB.
(*b*) By similar procedure, SNR = 34.9 dB. We observe that the SNR deteriorates as the operating current is lowered.

Flicker Noise

Flicker noise, also called $1/f$ *noise,* or *contact noise,* is present in all active as well as in some passive devices and has various origins, depending on device type. In active devices it is due to traps, which, when current flows, capture and release charge carriers randomly, thus causing random fluctuations in the current itself. In BJTs these traps are associated with contamination and crystal defects at the base-emitter junction. In MOSFETs they are associated with extra electron energy states at the boundary between silicon and silicon dioxide. Among active devices, MOSFETs suffer the most, and this can be a source of concern in low-noise MOS applications.

Flicker noise is always associated with a dc current, and its power density is of the type

$$i_n^2 = K\frac{I^a}{f} \tag{7.17}$$

where K is a device constant, I is the dc current, and a is another device constant in the range $\frac{1}{2}$ to 2.

Flicker noise is also found in some passive devices, such as carbon composition resistors, in which case it is called *excess noise* because it appears in addition to the thermal noise already there. However, while thermal noise is also present without a dc current, flicker noise requires a dc current in order to exist. Resistors of the wire-wound type are the quietest in terms of $1/f$ noise, while the carbon composition types can be noisier by as much as an order of magnitude, depending on operating conditions. Carbon-film and metal-film types fall in between. However, if the application requires that a given resistor carry a fairly small current, thermal noise will predominate and it will make little difference which resistor type one uses.

Avalanche Noise

This form of noise is found in *pn* junctions operated in the reverse breakdown mode. Avalanche breakdown occurs when electrons, under the influence of the strong electric field inside the space-charge layer, acquire enough kinetic energy to create additional electron-hole pairs by collision against the atoms of the crystal lattice. These additional pairs can, in turn, create other pairs in avalanche fashion. The resulting current consists of randomly distributed noise spikes flowing through the reverse biased junction. Like shot noise, avalanche noise requires current flow. However, avalanche noise is usually much more intense than shot noise, making Zener diodes notoriously noisy. This is one of the reasons why voltage references of the bandgap type are preferable to Zener-diode references.

Noise Models of Semiconductor Devices

We now wish to examine the noise mechanisms intervening in semiconductor devices so as to help the reader develop a basic feel for the noise characteristics of op amps. As a rule, we model each device with a noiseless version, but equipped with suitable noise sources e_n and i_n, in the manner already seen for the resistor in Fig. 7.6. The results are summarized in Table 7.1.

TABLE 7.1
Noise models and noise power densities for semiconductor devices (the devices in the models are assumed *noiseless* and noise is accounted for by suitable noise sources, as shown)

$$e_n^2 = 4kTr_S$$

$$i_n^2 = 2qI_D + K\frac{I_D^a}{f}$$

$$e_n^2 = 4kT\left(r_b + \frac{1}{2g_m}\right)$$

$$i_n^2 = 2q\left(I_B + K_1\frac{I_B^a}{f} + \frac{I_C}{|\beta_0(if)|^2}\right)$$

$$e_n^2 = 4kT\left(\frac{2}{3g_m} + K_2\frac{I_D^a/g_m^2}{f}\right)$$

$$i_n^2 = 2qI_G + \left(\frac{2\pi fC_{gs}}{g_m}\right)^2\left(4kT\frac{2}{3}g_m + K_3\frac{I_D^a}{f}\right)$$

$$e_n^2 = 4kT\frac{2}{3}\frac{1}{g_m} + K_4\frac{1}{WLf}$$

$$i_n^2 = 2qI_G$$

A *pn* junction generates both shot and flicker noise, so we model it with a noiseless diode but having a parallel noise current i_n, as depicted in Table 7.1. Also shown is a series noise voltage e_n modeling the thermal noise of the diode's bulk resistance r_S. The power densities of these source are[8]

$$i_n^2 = 2qI_D + \frac{KI_D^a}{f} \qquad (7.18a)$$

$$e_n^2 = 4kTr_S \qquad (7.18b)$$

The noise power densities of BJTs are[8]

$$e_n^2 = 4kT\left(r_b + \frac{1}{2g_m}\right) \tag{7.19a}$$

$$i_n^2 = 2q\left(I_B + K_1\frac{I_B^a}{f} + \frac{I_C}{|\beta(jf)|^2}\right) \tag{7.19b}$$

where r_b is the intrinsic base resistance, I_B and I_C are the dc base and collector currents, $g_m = qI_C/(kT)$ is the transconductance, K_1 and a are appropriate device constants, and $\beta(jf)$ is the forward current gain, which decreases at high frequencies.

In the expression for e_n^2, the first term represents thermal noise from r_b, and the second term represents the effect of collector-current shot noise referred to the input. In the expression for i_n^2, the first two terms represent base-current shot and flicker noise, and the last term represents collector-current shot noise reflected to the input.

To achieve a high β, the base region of a BJT is doped lightly and fabricated very thin. This, however, increases the intrinsic base resistance r_b. Moreover, the transconductance g_m and the base current I_B are directly proportional to I_C. Thus, what works to minimize voltage noise (low r_b and high I_C) is the opposite of what is good for low current noise (high β and low I_C). This represents a fundamental tradeoff in bipolar-op-amp design.

The noise power densities of JFETs are[8]

$$e_n^2 = 4kT\left(\frac{2}{3g_m} + K_2\frac{I_D^a/g_m^2}{f}\right) \tag{7.20a}$$

$$i_n^2 = 2qI_G + \left(\frac{2\pi f C_{gs}}{g_m}\right)^2\left(4kT\frac{2}{3}g_m + K_3\frac{I_D^a}{f}\right) \tag{7.20b}$$

where g_m is the transconductance; I_D is the dc drain current; I_G is the gate leakage current; K_2, K_3, and a are appropriate device constants; and C_{gs} is the gate-to-source capacitance.

In the expression for e_n^2, the first term represents thermal noise in the channel, and the second represents drain-current flicker noise. At room temperature and at moderate frequencies, all terms in the expression for i_n^2 are negligible, making JFETs virtually free of input current noise. Recall, however, that gate leakage increases very rapidly with temperature, so i_n^2 may no longer be neglected at higher temperatures.

Compared to BJTs, FETs have notoriously low g_m values, indicating that FET-input op amps tend to exhibit higher voltage noise than BJT-input types for similar operating conditions. Moreover, e_n^2 in the JFET contains flicker noise. These disadvantages are offset by better current noise performance, at least near room temperature.

The noise power densities of MOSFETs are[8]

$$e_n^2 = 4kT\frac{2}{3g_m} + K_4\frac{1}{WLf} \tag{7.21a}$$

$$i_n^2 = 2qI_G \tag{7.21b}$$

where g_m is the transconductance, K_4 is a device constant, and W and L are the channel width and length. As in the JFET case, i_n^2 is negligible at room temperature, but increases with temperature.

In the expression for e_n^2, the first term represents thermal noise from the channel resistance and the second represents flicker noise. It is the latter that is of most concern in MOSFET-input op amps. Flicker noise is inversely proportional to the transistor area $W \times L$, so this type of noise is reduced by using input-stage transistors with large geometries. As discussed in Chapter 5, when large geometries are combined with common-centroid layout techniques, the input offset voltage and offset drift characteristics are also improved significantly.

Noise Modeling in PSpice

When performing noise analysis, SPICE calculates the thermal-noise density for each resistor in the circuit, as well as the shot-noise and flicker-noise densities for each diode and transistor. When using op amp macromodels, the need arises for noise sources with spectral densities of the type of Fig. 7.2. We shall synthesize these sources[2] by exploiting the fact that SPICE calculates the noise current of a diode according to

$$i_d^2 = \mathrm{KF}\frac{I_D^{\mathrm{AF}}}{f} + 2qI_D = 2qI_D\left(\frac{\mathrm{KF} \times I_D^{\mathrm{AF}-1}/2q}{f} + 1\right)$$

where I_D is the diode bias current, q the electron charge, and KF and AF are parameters that can be specified by the user. This is a power density with white-noise floor $i_w^2 = 2qI_D$ and corner frequency $f_c = \mathrm{KF} \times I_D^{\mathrm{AF}-1}/2q$. If we let AF = 1 for mathematical convenience, then the required I_D and KF for given i_w^2 and f_c are

$$I_D = i_w^2/2q \qquad \mathrm{KF} = 2qf_c$$

Once we have a source of current noise, we can readily convert it to a source of voltage noise via a CCVS.

EXAMPLE 7.6. Verify Example 7.1 using PSpice.

Solution. We need to create a source e_n with $e_{nw} = 20\ \mathrm{nV}/\sqrt{\mathrm{Hz}}$ and $f_{ce} = 200$ Hz. First we create a noise current source with $i_w = 1\ \mathrm{pA}/\sqrt{\mathrm{Hz}}$ and $f_c = 200$ Hz, then we use an H-type source of value 20 nV/pA to convert to e_n. As shown in Fig. 7.7a, we bias the diode with $I_D = (1 \times 10^{-12})^2/(2 \times 1.602 \times 10^{-19}) = 3.12\ \mu$A, and we impose $\mathrm{KF} = 2 \times 1.602 \times 10^{-19} \times 200 = 6.41 \times 10^{-17}$ A. The 1-GF capacitor couples the ac noise current generated by the diode to the CCVS denoted as H. The PSpice model for the diode is

```
.model Dnoise D(KF=6.41E-17,AF=1)
```

The results of the simulation are shown in Fig. 7.7b. Using PSpice's cursor facility to measure specific values, we find that for $0.1\ \mathrm{Hz} \le f \le 100$ Hz, $E_n \cong 0.77\ \mu$V; for $20\ \mathrm{Hz} \le f \le 20$ kHz, $E_n \cong 3\ \mu$V; and for $0.1\ \mathrm{Hz} \le f \le 1$ MHz, $E_n = 20\ \mu$V. This corroborates the results of the hand calculations of Example 7.1.

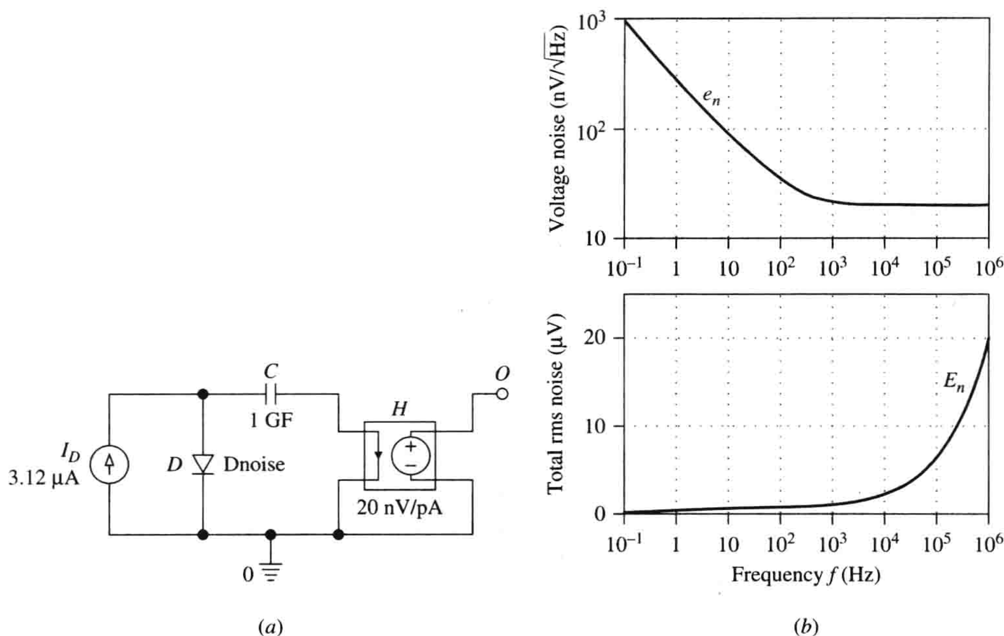


FIGURE 7.7
(*a*) Using a diode to synthesize a PSpice source of voltage noise. (*b*) Its spectral density e_n (top) is plotted as `V(ONOISE)`, and the total rms noise E_n (bottom) as `SQRT(S(V(ONOISE)*V(ONOISE)))`.

7.4 OP AMP NOISE

Op amp noise is characterized by three equivalent noise sources: a voltage source with spectral density e_n, and two current sources with densities i_{np} and i_{nn}. As shown in Fig. 7.8, a practical op amp can be regarded as a noiseless op amp equipped with these sources at the input. This model is similar to that used to account for the input offset voltage V_{OS} and the input bias currents I_P and I_N. This is not surprising

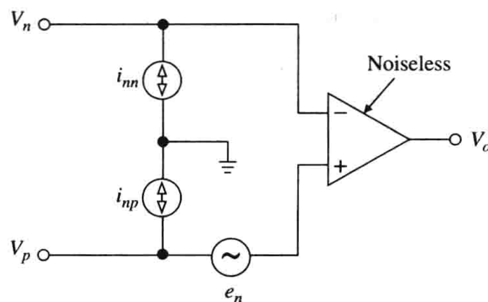


FIGURE 7.8
Op amp noise model.

since these parameters are themselves special forms of noise, namely, dc noise. Note, however, that the magnitudes and directions of $e_n(t)$, $i_{np}(t)$, and $i_{nn}(t)$ are constantly changing due to the random nature of noise and that noise terms must be added up in rms rather than algebraic fashion.

Noise densities are given in the data sheets and have the typical forms of Fig. 7.2. For devices with symmetric input circuitry, such as voltage-mode op amps (VFAs), i_{np} and i_{nn} are given as a single density i_n, even though i_{np} and i_{nn} are uncorrelated. To avoid losing track of their identities, we shall use separate symbols until the end of our calculations, when we shall substitute i_n for both i_{np} and i_{nn}. For current-feedback amplifiers (CFAs), the inputs are asymmetric due to the presence of the input buffer. Consequently, i_{np} and i_{nn} are different and are graphed separately.

Just as in precision dc applications it is important to know the dc output error E_O caused by V_{OS}, I_P, and I_N, in low-noise applications it is of interest to know the total rms output noise E_{no}. Once E_{no} is known, we can refer it back to the input and compare it against the useful signal to determine the signal-to-noise ratio SNR and, hence, the ultimate resolution of the circuit. We shall illustrate for the familiar resistive-feedback circuit of Fig. 7.9*a*, which forms the basis of the inverting and noninverting amplifiers, the difference and summing amplifiers, and a variety of others. It is important to keep in mind that the resistances shown in the diagram must include also the external source resistances, if any. For instance, if we lift node A off ground and drive it with a source v_S having internal resistance R_s, then we must replace R_1 with the sum $R_s + R_1$ in our calculations.

To analyze the circuit, we redraw it as in Fig. 7.9*b* with all pertinent noise sources in place, including the thermal noise sources of the resistors. As we know, resistor noise can be modeled with either a series voltage source or a parallel current source. The reason for choosing the latter will become apparent shortly.

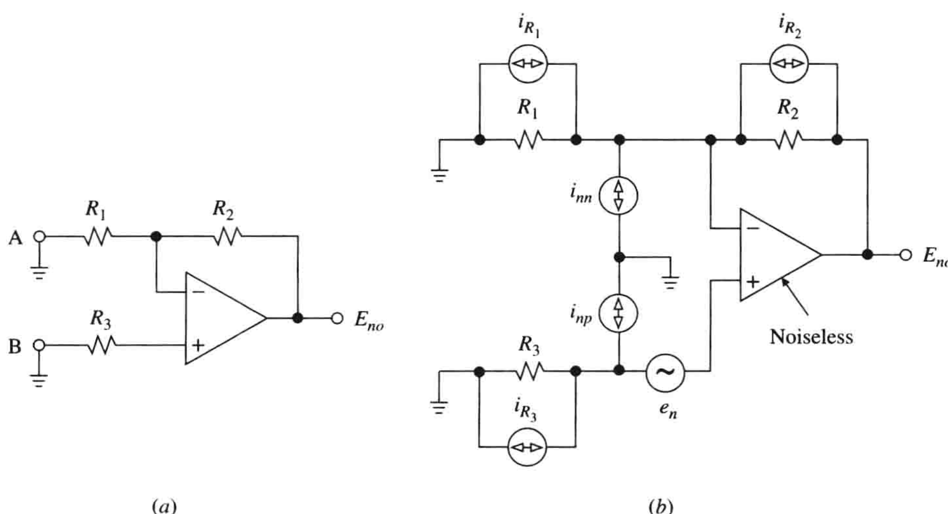


FIGURE 7.9
Resistive-feedback op amp circuit and its noise model.

Overall Input Spectral Density

The first task is to find the overall spectral density e_{ni} referred to the input of the op amp. We can apply the superposition principle as when we calculate the overall input error E_I due to V_{OS}, I_P, and I_N, except that now the individual terms must be added up in rms fashion. Thus, the noise voltage e_n contributes the term e_n^2. The noise currents i_{np} and i_{R_3} are flowing through R_3, so their combined contribution is, by Eq. (7.15), $(R_3 i_{np})^2 + (R_3 i_{R_3})^2 = R_3^2 i_{np}^2 + 4kTR_3$. Following the line of reasoning that led to Eq. (5.10), we can say that the noise currents i_{nn}, i_{R_1}, and i_{R_2} conspire with the combination $R_1//R_2$ to contribute $(R_1 \| R_2)^2 (i_{nn}^2 + i_{R_1}^2 + i_{R_2}^2) = (R_1 \| R_2)^2 i_{nn}^2 + 4kT(R_1 \| R_2)$. Combining all terms gives the overall input spectral density

$$e_{ni}^2 = e_n^2 + R_3^2 i_{np}^2 + (R_1 \| R_2)^2 i_{nn}^2 + 4kT[R_3 + (R_1 \| R_2)] \qquad (7.22)$$

For op amps with symmetric inputs and uncorrelated noise currents, we have $i_{np} = i_{nn} = i_n$, where i_n is the noise current density given in the data sheets.

To gain better insight into the relative weights of the various terms, consider the special but familiar case in which $R_3 = R_1//R_2$. Under this constraint, Eq. (7.22) simplifies as

$$e_{ni}^2 = e_n^2 + 2R^2 i_n^2 + 8kTR \qquad (7.23a)$$

$$R = R_1//R_2 = R_3 \qquad (7.23b)$$

Figure 7.10 shows e_{ni} as well as its three individual components as a function of R. While the voltage term e_n is independent of R, the current term $\sqrt{2}Ri_n$ increases with R at the rate of 1 dec/dec, and the thermal term $\sqrt{8kTR}$ increases at the rate of 0.5 dec/dec.

We observe that for R sufficiently small, voltage noise dominates. In the limit $R \to 0$ we get $e_{ni} \to e_n$, so e_n is aptly called the *short-circuit* noise: this is the noise produced by the internal components of the op amp, regardless of the external circuitry. For R sufficiently large, current noise dominates. In the limit $R \to \infty$ we get $e_{ni} \to \sqrt{2}Ri_n$, so i_n is aptly called the *open-circuit* noise. This form of noise stems from input-bias-current flow through the external resistors. For intermediate values of R, thermal noise may also come into play, depending on the relative magnitudes

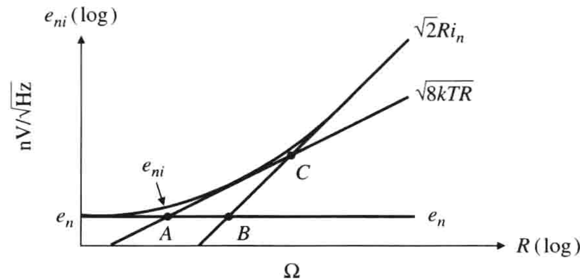


FIGURE 7.10
Op amp input spectral noise e_{ni} as a function of R of Eq. (7.23b).

of the other two terms. In the example pictured, point A is where thermal noise overtakes voltage noise, point B where current noise overtakes voltage noise, and point C where current noise overtakes thermal noise. The relative positions of A, B, and C vary from one op amp to another, and can be used to compare different devices.

We note that while it is desirable to install a dummy resistance $R_3 = R_1 \parallel R_2$ in order to provide bias-current compensation, in terms of noise it is preferable to have $R_3 = 0$ since this resistor only contributes additional noise. When the presence of R_3 is mandatory, the corresponding thermal noise can be filtered out by connecting a suitably large capacitance in parallel with R_3. This will also suppress any external noise that might be accidentally injected into the noninverting input pin.

Rms Output Noise

Like offsets and drift, e_{ni} is amplified by the noise gain of the circuit. This gain is not necessarily the same as the signal gain, so we shall denote *signal gain* as $A_s(jf)$ and *noise gain* as $A_n(jf)$ to avoid confusion. Recall that the dc value of $A_n(jf)$ is $A_{n0} = 1/\beta = 1 + R_2/R_1$. Moreover, for a constant-GBP op amp, the closed-loop bandwidth of $A_n(jf)$ is $f_B = \beta f_t = f_t/(1 + R_2/R_1)$, where f_t is the unity-gain frequency of the op amp. The output spectral density can thus be expressed as

$$e_{no} = \frac{1 + R_2/R_1}{\sqrt{1 + (f/f_B)^2}} e_{ni} \tag{7.24}$$

Noise is observed or measured over a finite time interval T_{obs}. The *total rms output noise* is found by integrating e_{no}^2 from $f_L = 1/T_{\text{obs}}$ to $f_H = \infty$. Using Eqs. (7.9), (7.11*b*), (7.14*b*), and (7.22) we get

$$\begin{aligned} E_{no} = \left[1 + \frac{R_2}{R_1}\right] \times \bigg[& e_{nw}^2 \left(f_{ce} \ln \frac{f_B}{f_L} + 1.57 f_B - f_L\right) \\ & + R_3^2 i_{npw}^2 \left(f_{cip} \ln \frac{f_B}{f_L} + 1.57 f_B - f_L\right) \\ & + (R_1 \parallel R_2)^2 i_{nnw}^2 \left(f_{cin} \ln \frac{f_B}{f_L} + 1.57 f_B - f_L\right) \\ & + 4kT(R_3 + R_1 \parallel R_2)(1.57 f_B - f_L)\bigg]^{1/2} \end{aligned} \tag{7.25}$$

This expression indicates the considerations in low-noise design: (*a*) select op amps with low-noise floors e_{nw} and i_{nw} as well as low corner frequencies f_{ce} and f_{ci}; (*b*) keep the external resistances sufficiently small to make current noise and thermal noise negligible compared to voltage noise (if possible, make $R_3 = 0$); (*c*) limit the noise-gain bandwidth to the strict minimum required.

The industry standard OP77 op amp has been specifically designed for low-noise applications. Its characteristics are $f_t = 8$ MHz, $e_{nw} = 3$ nV/$\sqrt{\text{Hz}}$, $f_{ce} = 2.7$ Hz, $i_{nw} = 0.4$ pA/$\sqrt{\text{Hz}}$, and $f_{ci} = 140$ Hz.

EXAMPLE 7.7. A 741 op amp is configured as an inverting amplifier with $R_1 = 100\ \text{k}\Omega$, $R_2 = 200\ \text{k}\Omega$, and $R_3 = 68\ \text{k}\Omega$. (*a*) Assuming $e_{nw} = 20\ \text{nV}/\sqrt{\text{Hz}}$, $f_{ce} = 200$ Hz, $i_{nw} = 0.5\ \text{pA}/\sqrt{\text{Hz}}$, and $f_{ci} = 2$ kHz, find the total output noise above 0.1 Hz, both rms and peak-to-peak. (*b*) Verify with PSpice.

Solution.

(*a*) We have $R_1 \parallel R_2 = 100 \parallel 200 \cong 67\ \text{k}\Omega$, $A_{n0} = 1 + R_2/R_1 = 3$ V/V, and $f_B = 10^6/3 = 333$ kHz. The noise voltage component is $E_{noe} = 3 \times 20 \times 10^{-9}[200 \ln(333 \times 10^3/0.1) + 1.57 \times 333 \times 10^3 - 0.1]^{1/2} = 43.5\ \mu\text{V}$. The current noise component is $E_{noi} = 3[(68 \times 10^3)^2 + (67 \times 10^3)^2]^{1/2} \times 0.5 \times 10^{-12} \times [2 \times 10^3 \ln(333 \times 10^4) + 523 \times 10^3]^{1/2} = 106.5\ \mu\text{V}$. The thermal noise component is $E_{noR} = 3[1.65 \times 10^{-20}(68 + 67) \times 10^3 \times 523 \times 10^3]^{1/2} = 102.4\ \mu\text{V}$. Finally,

$$E_{no} = \sqrt{E_{noe}^2 + E_{noi}^2 + E_{noR}^2} = \sqrt{43.5^2 + 106.5^2 + 102.4^2} = 154\ \mu\text{V (rms)}$$

or $6.6 \times 154 = 1.02$ mV (peak-to-peak).

(*b*) The 741 macromodel available with the student version of PSpice does not model noise properly. Consequently, we use a Laplace block to simulate a noiseless op amp having the same open-loop response as the 741, along with three noise sources of the type of Fig 7.7*a* to simulate e_n, i_{nn}, and i_{np} (the sources must be separate in order to ensure statistical independence). The circuit, shown in Fig. 7.11, uses diodes with the same PSpice model of Fig. 7.7*a*,

```
.model Dnoise D(KF=6.41E-17,AF=1)
```

The plot of Fig. 7.12*b* confirms the hand calculation of $E_{no} = 154\ \mu$V. In hindsight, we could have estimated E_{no} using just the pink-noise tangent principle! Indeed, Fig. 7.12*a* indicates that most of the noise comes from the noise floor of $210\ \text{nV}/\sqrt{\text{Hz}}$ preceding the pole frequency of 333 kHz, so $E_{no} \cong (210\ \text{nV}) \times (1.57 \times 333\ \text{kHz})^{1/2} = 152\ \mu\text{V}$ (rms), quite close to the actual vale of 154 μV (rms)!

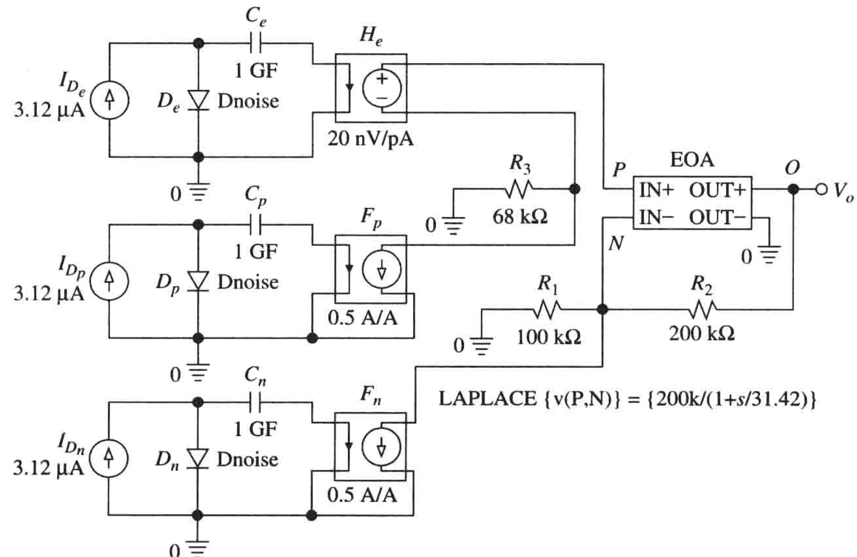


FIGURE 7.11
PSpice circuit to measure the noise of a 741-like op amp circuit.

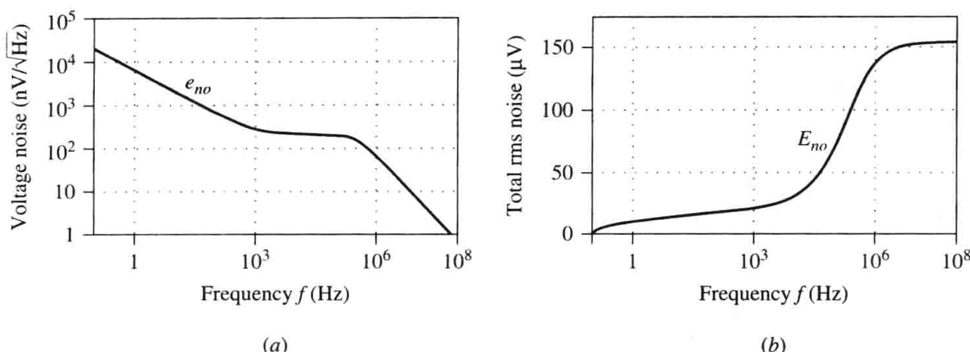


FIGURE 7.12
(*a*) Plotting the output spectral density e_{no} of the circuit of Fig. 7.11 as `V(ONOISE)`, and (*b*) plotting the total rms output noise E_{no} as `SQRT(S(V(ONOISE)*V(ONOISE)))`.

In terms of noise, the circuit of Example 7.7 is poorly designed because E_{noi} and E_{noR} far exceed E_{noe}. This can be improved by scaling down all resistances. A good rule of thumb is to impose $E_{noi}^2 + E_{noR}^2 \le E_{noe}^2/3^2$, since this raises E_{no} only by about 5%, or less, above E_{noe}.

EXAMPLE 7.8. Scale the resistances of the circuit of Example 7.7 so that $E_{no} = 50\ \mu$V.

Solution. We want $E_{noi}^2 + E_{noR}^2 = E_{no}^2 - E_{noe}^2 = 50^2 - 43.5^2 = (24.6\ \mu\text{V})^2$. Letting $R = R_3 + R_1 \parallel R_2$, we have $E_{noi}^2 = 3^2 \times R^2(0.5 \times 10^{-12})^2 \times [2 \times 10^3 \ln(333 \times 10^4) + 523 \times 10^3] = 1.24 \times 10^{-18} R^2$, andx $E_{noR}^2 = 3^2 \times 1.65 \times 10^{-20} \times R \times 523 \times 10^3 = 7.77 \times 10^{-14} R$. We want $1.24 \times 10^{-18} R^2 + 7.77 \times 10^{-14} R = (24.6\ \mu\text{V})^2$, which gives $R = 7$ kΩ. Thus, $R_3 = R/2 = 3.5$ kΩ, and $1/R_1 + 1/R_2 = 1/(3.5\ \text{k}\Omega)$. Since $R_2 = 2R_1$, this yields $R_1 = 5.25$ kΩ and $R_2 = 10.5$ kΩ.

Signal-to-Noise Ratio

Dividing E_{no} by the dc signal gain $|A_{s0}|$ yields the *total rms input noise,*

$$E_{ni} = \frac{E_{no}}{|A_{s0}|} \tag{7.26}$$

We again stress that the signal gain A_s may be different from the noise gain A_n, the inverting amplifier being a familiar example. Knowing E_{ni} allows us to find the *input signal-to-noise ratio,*

$$\text{SNR} = 20 \log_{10} \frac{V_{i(\text{rms})}}{E_{ni}} \tag{7.27}$$

where $V_{i(\text{rms})}$ is the rms value of the input voltage. The SNR establishes the ultimate resolution of the circuit. For an amplifier of the transimpedance type, the total rms input noise is $I_{ni} = E_{no}/|R_{s0}|$, where $|R_{s0}|$ is the dc transimpedance signal gain. Then, $\text{SNR} = 20 \log_{10}(I_{i(\text{rms})}/I_{ni})$.

EXAMPLE 7.9. Find the SNR of the circuit of Example 7.7 if the input is an ac signal with a peak amplitude of 0.5 V.

Solution. Since $A_{s0} = -2$ V/V, we have $E_{ni} = 154/2 = 77\ \mu$V. Moreover, $V_{i(\text{rms})} = 0.5/\sqrt{2} = 0.354$ V. So, SNR $= 20\log_{10}[0.354/(77\times10^{-6})] = 73.2$ dB.

Noise in CFAs

The above equations apply also to CFAs.[9] As mentioned, the presence of the input buffer makes the inputs asymmetric, so i_{np} and i_{nn} are different. Moreover, since CFAs are wideband amplifiers, they generally tend to be noisier than conventional op amps.[10]

EXAMPLE 7.10. The data sheets of the CLC401 CFA (Comlinear) give $z_0 \cong 710$ kΩ, $f_b \cong 350$ kHz, $r_n \cong 50$ Ω, $e_{nw} \cong 2.4$ nV/$\sqrt{\text{Hz}}$, $f_{ce} \cong 50$ kHz, $i_{npw} \cong 3.8$ pA/$\sqrt{\text{Hz}}$, $f_{cip} \cong 100$ kHz, $i_{nnw} \cong 20$ pA/$\sqrt{\text{Hz}}$, and $f_{cin} \cong 100$ kHz. Find the total rms output noise above 0.1 Hz if the CFA is configured as a noninverting amplifier with $R_1 = 166.7$ Ω and $R_2 = 1.5$ kΩ, and is driven by a source with an internal resistance of 100 Ω.

Solution. Since $f_t = z_0 f_b/R_2 = 166$ MHz, we have $f_B = f_t/[1 + r_n/(R_1 \parallel R_2)] = 124$ MHz. Applying Eq. (7.25) gives $E_{no} = 10[(33.5\ \mu\text{V})^2 + (3.6\ \mu\text{V})^2 + (35.6\ \mu\text{V})^2 + (28.4\ \mu\text{V})^2]^{1/2} \cong 566\ \mu$V (rms), or $6.6 \times 566 \cong 3.7$ mV (peak-to-peak).

Noise Filtering

Since broadband noise increases with the square root of the noise-gain bandwidth, noise can be reduced through narrowbanding. The most common technique is to pass the signal through a simple *R-C* network with *R* small enough to avoid adding appreciably to the existing noise. This filter is susceptible to output loading, so we may want to buffer it with a voltage follower. However, this would add the noise of the follower, whose equivalent bandwidth NEB $= (\pi/2) f_t$ is quite wide.

The topology[11] of Fig. 7.13 places the op amp upstream of the *R-C* network so that the noise of the op amp itself is filtered. Moreover, *R* is placed within the feedback loop to reduce its effective value by $1 + T$ and thus reduce output loading

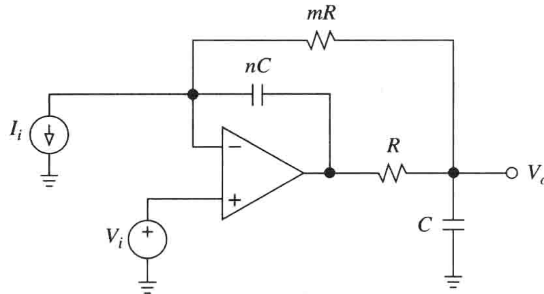


FIGURE 7.13
Low-pass noise filter. Input may be either a current or a voltage.

significantly. Even though T decreases with frequency, the presence of C helps maintain a low output impedance well into the upper frequency range. The purpose of mR and nC is to provide frequency compensation, an issue addressed in Section 8.4. Suffice it to say here that the circuit exhibits a good tolerance to capacitive loads.

The circuit lends itself to filtering both voltages and currents. It can be shown (see Problem 7.31) that

$$V_o = H_{\text{LP}} mRI_i + (H_{\text{LP}} + H_{\text{BP}})V_i \qquad (7.28)$$

$$f_0 = \frac{1}{2\pi\sqrt{mn}RC} \qquad Q = \frac{\sqrt{m/n}}{m+1} \qquad (7.29)$$

where H_{LP} and H_{BP} are the standard second-order low-pass and band-pass functions defined in Section 3.3. This filter finds application in voltage-reference and photodiode-amplifier noise reduction.

7.5 NOISE IN PHOTODIODE AMPLIFIERS

An area in which noise is of concern is low-level signal detection, such as instrumentation applications and high-sensitivity *I-V* conversion. In particular, photodiode amplifiers have been at the center of considerable attention,[12] so we examine this class of amplifiers in some detail.

The photodiode of Fig. 7.14*a* responds to incident light with a current i_S that the op amp subsequently converts to a voltage v_O. For a realistic analysis we use the model of Fig. 7.14*b*, where R_1 and C_1 represent the combined resistance and capacitance toward ground of the diode and the inverting-input pin of the op amp, and C_2 represents the stray capacitance of R_2. With careful printed-circuit board layout, C_2 can be kept in the range of 1 pF or less. Usually $C_1 \gg C_2$ and $R_1 \gg R_2$.

We are interested in the signal gain $A_s = V_o/I_s$ as well as the noise gain $A_n = e_{no}/e_{ni}$. To this end, we need to find the feedback factor $\beta = Z_1/(Z_1 + Z_2)$,

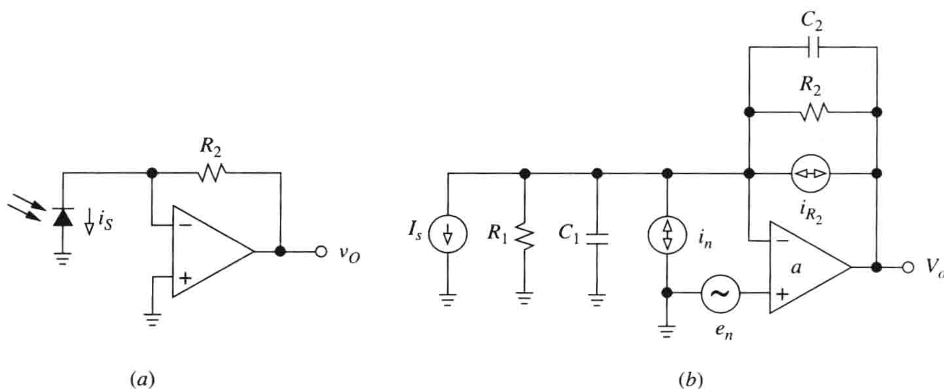


FIGURE 7.14
Photodiode amplifier and its noise model.

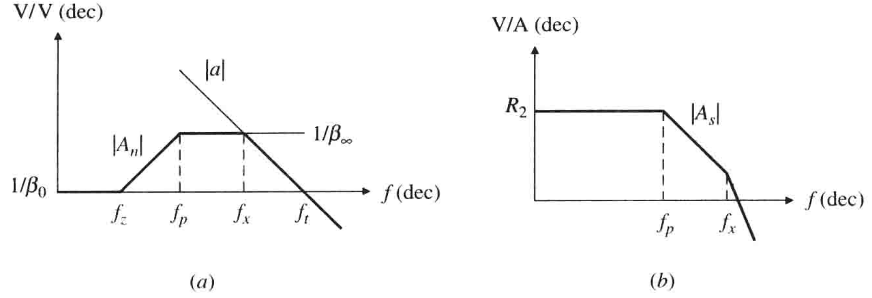


FIGURE 7.15
Noise gain A_n and signal gain A_s for the photodiode amplifier.

$Z_1 = R_1 \parallel [1/(j2\pi f C_1)]$, $Z_2 = R_2 \parallel [1/(j2\pi f C_2)]$. Expanding gives

$$\frac{1}{\beta} = \left(1 + \frac{R_2}{R_1}\right) \frac{1 + jf/f_z}{1 + jf/f_p} \tag{7.30a}$$

$$f_z = \frac{1}{2\pi (R_1 \parallel R_2)(C_1 + C_2)} \qquad f_p = \frac{1}{2\pi R_2 C_2} \tag{7.30b}$$

The $1/\beta$ function has the low-frequency asymptote $1/\beta_0 = 1 + R_2/R_1$, the high-frequency asymptote $1/\beta_\infty = 1 + C_1/C_2$, and two breakpoints at f_z and f_p. As shown in Fig. 7.15*a*, the crossover frequency is $f_x = \beta_\infty f_t$, so the noise gain is $A_n = (1/\beta)/(1 + jf/f_x)$, or

$$A_n = \left(1 + \frac{R_2}{R_1}\right) \frac{1 + jf/f_z}{(1 + jf/f_p)(1 + jf/f_x)} \tag{7.31}$$

We also observe that for $a \to \infty$ we have $A_{s(\text{ideal})} = R_2/(1 + jf/f_p)$, so the signal gain is

$$A_s = \frac{R_2}{(1 + jf/f_p)(1 + jf/f_x)} \tag{7.32}$$

and is shown in Fig. 7.15*b*. With $C_1 \gg C_2$, the noise-gain curve exhibits significant peaking, a notorious feature of photodiode amplifiers. This can be reduced by adding a capacitance in parallel with R_2; however, this also reduces the signal-gain bandwidth f_p.

EXAMPLE 7.11. In the circuit[12] of Fig. 7.14 let the op amp be the OPA627 JFET-input op amp, for which $f_t = 16$ MHz, $e_{nw} = 4.5$ nV/$\sqrt{\text{Hz}}$, $f_{ce} = 100$ Hz, and $I_B = 1$ pA. Estimate the total output noise E_{no} above 0.01 Hz if $R_1 = 100$ GΩ, $C_1 = 45$ pF, $R_2 = 10$ MΩ, and $C_2 = 0.5$ pF.

Solution. With the above data we have $1/\beta_0 \cong 1$ V/V, $1/\beta_\infty = 91$ V/V, $f_z = 350$ Hz, $f_p = 31.8$ kHz, and $f_x = 176$ kHz. Moreover, by Eqs. (7.15*b*) and (7.16), $i_{R_2} =$ 40.6 fA/$\sqrt{\text{Hz}}$ and $i_n = 0.566$ fA/$\sqrt{\text{Hz}}$. We observe that the noise gain for e_n is A_n, whereas the noise gains for i_n and i_{R_2} coincide with the signal gain A_s. The output densities, obtained as $e_{noe} = |A_n| e_n$, $e_{noi} = |A_s| i_n$, and $e_{noR} = |A_s| i_{R_2}$, are plotted in Fig. 7.16.

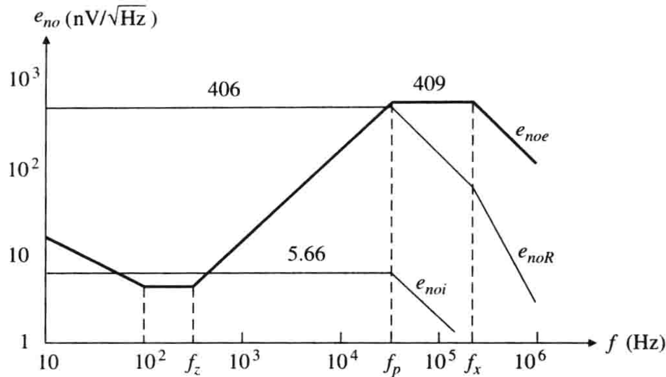


FIGURE 7.16
Output spectral densities of the photodiode amplifier of Example 7.11.

The pink-noise tangent principle reveals that the dominant components are the voltage noise e_{noe} in the vicinity of f_x, and the thermal noise e_{noR} in the vicinity of f_p. Current noise is negligible because we are using a JFET-input op amp. Thus, $E_{noe} \cong (1/\beta_\infty)e_n\sqrt{(\pi/2)f_x - f_p} = 91 \times 4.5 \times 10^{-9}\sqrt{(1.57 \times 176 - 31.8)10^3} = 202\ \mu\text{V}$ (rms), and $E_{noR} \cong R_2 i_{R_2} \times \sqrt{(\pi/2)f_p} \cong 91\ \mu\text{V}$. Finally, $E_{no} \cong \sqrt{202^2 + 91^2} = 222\ \mu\text{V}$ (rms). A PSpice simulation (see Problem 7.35) gives $E_{no} = 230\ \mu\text{V}$ (rms), indicating that our hand-calculation approximations are quite reasonable.

Noise Filtering

The modified photodiode amplifier of Fig. 7.17 incorporates the current-filtering option of Fig. 7.13 to reduce noise. In choosing the filter cutoff frequency f_0, we must be careful that the signal-gain bandwidth is not reduced unnecessarily. Moreover, the optimum value of Q is the result of a compromise between noise and response characteristics such as peaking and ringing. A reasonable approach is to start with

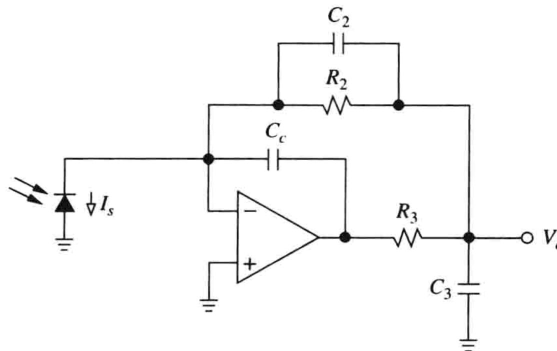


FIGURE 7.17
Photodiode amplifier with noise filtering.

$C_c = C_2$ and $R_3C_3 = R_2C_c$, so that $m = 1/n$ and $Q \cong 1$ for $m \gg 1$. Then we fine-tune C_c and R_3 for a best compromise between noise and response characteristics.

EXAMPLE 7.12. Assuming the parameters of Example 7.11, find suitable values for C_c, R_3, and C_3 in the circuit of Fig. 7.17.

Solution. Let $C_c = C_2 = 0.5$ pF. Pick $C_3 = 10$ nF as a convenient value. Then, $R_3 = R_2C_c/C_3 = 500\ \Omega$.

PSpice simulations for different values of R_3 give a good compromise for $R_3 = 1\ \text{k}\Omega$, which results in a signal-gain bandwidth of about 24 kHz and $E_{no} \cong 80\ \mu\text{V}$ (rms). Thus, filtering has reduced noise to about one third of the original value of 230 μV (rms). When the circuit is tried out in the lab, empirical tuning is necessary because of parasitics not accounted for by our PSpice model.

T-Feedback Photodiode Amplifiers

As we know, the use of a *T*-network makes it possible to achieve extremely high sensitivities using moderately high resistances. To assess its impact on dc as well as noise, we use the model of Fig. 7.18. The *T*-network is usually implemented with $R_3 \parallel R_4 \ll R_2$, so R_2 is raised to the equivalent value $R_{\text{eq}} \cong (1 + R_4/R_3)R_2$, and $i_R^2 \cong i_{R_2}^2 = 4kT/R_2$. One can show (see Problem 7.38) that the noise and signal gains are now

$$A_n \cong \left(1 + \frac{R_2}{R_1}\right)\left(1 + \frac{R_4}{R_3}\right)\frac{1 + jf/f_z}{(1 + jf/f_p)(1 + jf/f_x)} \tag{7.33a}$$

$$A_s \cong \frac{(1 + R_4/R_3)R_2}{(1 + jf/f_p)(1 + jf/f_x)} \tag{7.33b}$$

$$f_z = \frac{1}{2\pi(R_1 \parallel R_2)(C_1 + C_2)} \qquad f_p = \frac{1}{2\pi(1 + R_4/R_3)R_2C_2} \tag{7.34}$$

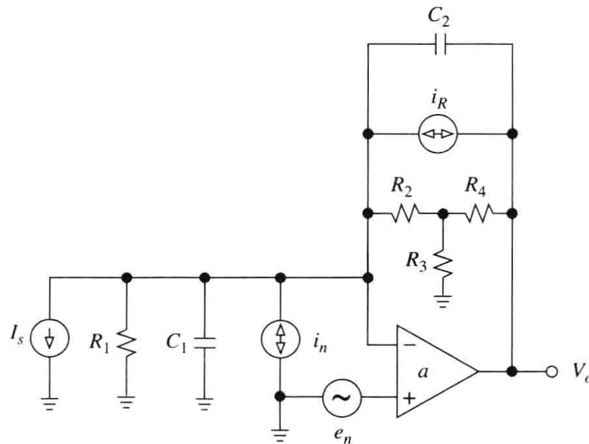


FIGURE 7.18
T-network photodiode amplifier.

indicating that the dc values of both gains are raised by a factor of $1 + R_4/R_3$. In particular, we observe that $E_{noR} \cong (1 + R_4/R_3) \times R_2 i_R \sqrt{\pi f_p/2} = [(1 + R_4/R_3)kT/C_2]^{1/2}$, indicating that thermal noise increases with the square root of the factor $1 + R_4/R_3$. Consequently, we must suitably limit this factor in order to avoid raising noise unnecessarily. As it turns out, the T-network option is worthwhile[12] when high-sensitivity amplifiers are used in connection with large-area photodiodes. The large capacitances of these devices cause enough noise-gain peaking to allow for thermal noise increase without jeopardizing the overall noise performance.

EXAMPLE 7.13. In the circuit[12] of Fig. 7.18 let the op amp be the OPA627 of Example 7.11, and let the diode be a large-area photodiode such that $C_1 = 2$ nF, everything else remaining the same. (*a*) Specify a T-network for a dc sensitivity of 1 V/nA. (*b*) Find the total rms output noise and the signal bandwidth.

Solution.

(*a*) We now have $1/\beta_0 \cong 1 + R_4/R_3$, $1/\beta_\infty = 1 + C_1/C_2 = 4000$ V/V and $f_x = \beta_\infty f_t = 4$ kHz. To avoid increasing voltage noise unnecessarily, impose $1/\beta_0 < 1/\beta_\infty$, or $1 + R_4/R_3 < 4000$. Then, $E_{noe} \cong (1/\beta_\infty) e_n \sqrt{\pi f_x/2} = 1.43$ mV. To avoid increasing thermal noise unnecessarily, impose $E_{noR} \leq E_{noe}/3$, or $[(1 + R_4/R_3)kT/C_2]^{1/2} \leq E_{noe}/3$. This yields $1 + R_4/R_3 \leq 27$ (< 4000). Then $R_2 = 10^9/27 = 37$ MΩ. Pick $R_2 = 36.5$ MΩ, $R_3 = 1.00$ kΩ, $R_4 = 26.7$ kΩ.

(*b*) The signal bandwidth is $f_B = f_p = 1/(2\pi \times 10^9 \times 0.5 \times 10^{-12}) = 318$ Hz. Moreover, $E_{noR} \cong 0.5$ mV, $E_{noi} = 10^9 \times 0.566 \times 10^{-15} \sqrt{1.57 \times 318} = 12.6$ μV, and $E_{no} \cong \sqrt{1.43^2 + 0.5^2} = 1.51$ mV (rms).

7.6 LOW-NOISE OP AMPS

As discussed in Section 7.4, the figures of merit in op amp noise performance are the white-noise floors e_{nw} and i_{nw}, and the corner frequencies f_{ce} and f_{ci}. The lower their values, the quieter the op amp. In wideband applications, usually only the white-noise floors are of concern; however, in instrumentation applications the corner frequencies may be crucial as well.

A well-documented[13] op amp that was specifically designed with the optimization of the above figures of merit in mind is the industry-standard OP27 of Fig. 7.19. The device incorporates a number of the precision features discussed in Chapter 5, namely, input current cancellation (Q_6), common-centroid layout (Q_{1A}/Q_{1B}-Q_{2A}/Q_{2B}), and on-chip V_{OS} trimming (R_1-R_2). It also includes an input-protection diode pair, as shown. The noise characteristics, graphed in Fig. 7.20*a* and *b*, are: $e_{nw} = 3$ nV/$\sqrt{\text{Hz}}$, $f_{ce} = 2.7$ Hz, $i_{nw} = 0.4$ pA/$\sqrt{\text{Hz}}$, and $f_{ci} = 140$ Hz. These characteristics are compared, in Fig. 7.20*c*, against those of the NE5533 low-noise audio op amp and the 741 general-purpose op amp. In the following discussion we shall use the OP27 as a vehicle for a discussion of basic low-noise design issues.

Except for programmable op amps, the user has no control over the noise characteristics; however, a basic understanding of how these characteristics originate will help in the device selection process. As with the input offset voltage and bias

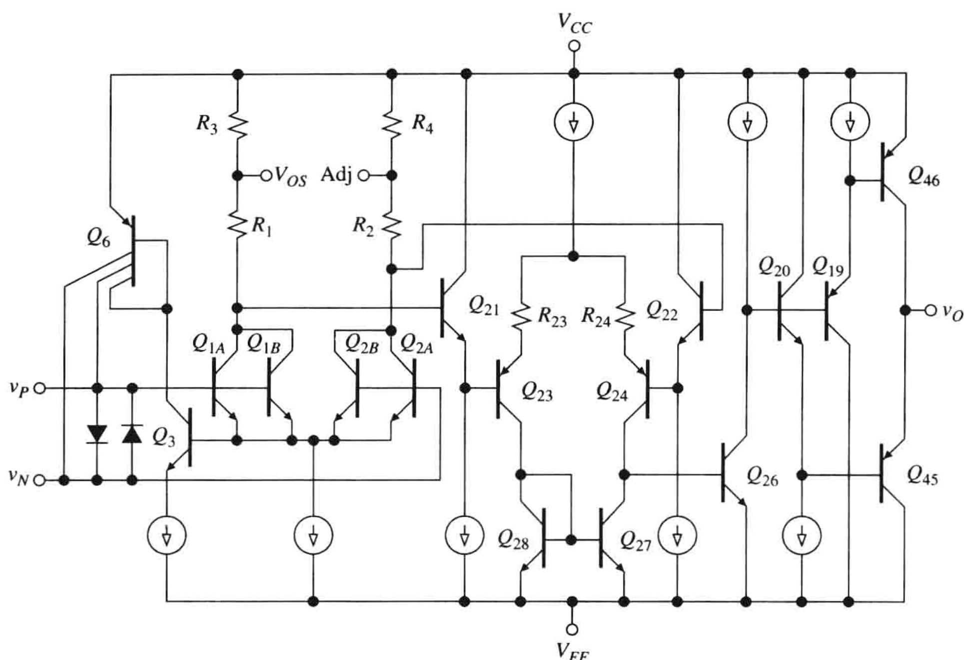


FIGURE 7.19
Simplified circuit diagram of the OP27 low-noise precision op amp. (Courtesy of Analog Devices.)

current, both voltage and current noise depend very heavily on the technology and operating conditions of the differential transistor pair of the input stage. Voltage noise is also affected by the load of the input pair and by the second stage. The noise produced by the subsequent stages is less critical because, as discussed in connection with Fig. 1.27, when reflected to the input, this noise gets divided by the gains of all preceding stages.

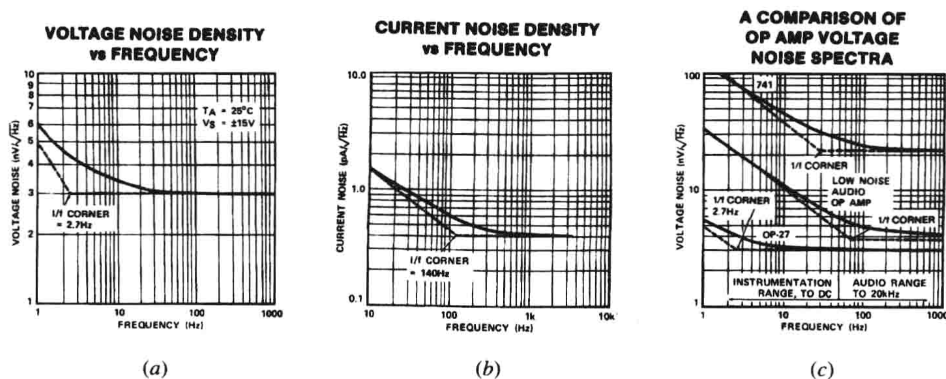


FIGURE 7.20
(*a*) Noise-voltage and (*b*) noise-current characteristics of the OP27/37 op amp. (*c*) Noise-voltage comparison of three popular op amps. (Courtesy of Analog Devices.)

Differential Input-Pair Noise

The noise contributed by the differential input pair can be minimized by proper choice of transistor type, geometry, and operating current. Consider BJT-input op amps first. Recall from Eq. (7.19*a*) that BJT voltage noise depends on the base-spreading resistance r_b and transconductance g_m. In the OP27 the differential-pair BJTs are realized in the *striped geometry* (long and narrow emitters surrounded by base contacts on both sides) to minimize r_b, and are biased at substantially higher-than-normal collector currents (120 μA per side) to increase g_m.[13] The increase in operating current, however, has an adverse effect on the input bias current I_B and the input noise current i_n. In the OP27 shown in Fig. 7.19, I_B is reduced by the current-cancellation technique. Noise densities, however, do not cancel but add up in rms fashion, so in current-cancellation schemes, i_{nw} is higher than the shot-noise value predicted by Eq. (7.19*b*).

When the application requires large external resistances, FET-input op amps offer a better alternative since their noise current levels are orders of magnitude lower than those of BJT-input devices, at least near room temperature. FETs, on the other hand, tend to exhibit higher voltage noise, mainly because they have lower g_ms than BJTs. As an example of a JFET-input op amp, the OPA827 has, at 1 kHz, $e_n = 4\ \text{nV}/\sqrt{\text{Hz}}$ and $i_n = 2.2\ \text{fA}/\sqrt{\text{Hz}}$.

In the case of MOSFETs, $1/f$ noise is also a critical factor. By Eq. (7.21*a*), the $1/f$ component can be reduced by using large-area devices. Moreover, the empirical observation that p-channel devices tend to display less $1/f$ noise than n-channel types indicates that, in general, the best noise performance in CMOS op amps is achieved by using p-channel input transistors. As an example of a CMOS op amp, the OPA320 has, at 1 kHz, $e_n = 8.5\ \text{nV}/\sqrt{\text{Hz}}$ and $i_n = 0.6\ \text{fA}/\sqrt{\text{Hz}}$ at 1 kHz. (A web search for low-noise JFET or CMOS op amps will give you a better idea of the staggering variety of available products.)

Input-Pair Load Noise

Another critical source of noise is the load of the differential input pair. In general-purpose op amps such as the 741, this load is implemented with a current-mirror active load to maximize gain. Active loads, however, are notoriously noisy since they amplify their own noise current. Once divided by the first-stage transconductance and converted to an equivalent input noise voltage, this component can degrade the noise characteristics significantly. In fact, in the 741, noise from the active load exceeds noise from the differential input pair itself.[8]

The OP27 avoids this problem by using a resistively loaded input stage,[13] as shown in Fig. 7.19. In CMOS op amps, the noise contribution from the active load, when reflected back to the input, is multiplied by the ratio of the g_m of the load to the g_m of the differential pair.[8] Thus, using a load with low g_m reduces this component significantly.

Second-Stage Noise

The last potentially critical contributor to e_n is the second stage, particularly when this is implemented with *pnp* transistors to provide level shifting as well as additional

gain (see Q_{23} and Q_{24} in Fig. 7.19). Being surface devices, *pnp* transistors suffer from large $1/f$ noise and poor β. Once this noise is reflected back to the input, it can increase f_{ce} significantly. The OP27 avoids this drawback by using emitter followers Q_{21} and Q_{22} (see again Fig. 7.19) to *isolate* the first stage from the *pnp* pair.[13]

Ultralow-Noise Op Amps

High-precision instrumentation often requires ultrahigh open-loop gains to achieve the desired degree of linearity, together with ultralow noise to ensure an adequate SNR. In these situations, considerations of cost and availability may justify the development of specialized circuits to meet the requirements.

Figure 7.21 shows an example of specialized op amp design whose dc specifications are compatible with high-precision transducer requirements and ac specifications are suitable for professional audio work.[14] The circuit uses the low-noise OP27 op amp with a differential front end to simultaneously increase the open-loop gain and reduce voltage noise. The front end consists of three parallel-connected MAT-02 low-noise dual BJTs operating at moderately high collector currents (1 mA per transistor). The parallel arrangement reduces the base spreading resistance of the composite device by $\sqrt{3}$, while the high collector current increases g_m. This yields an equivalent input noise voltage with $e_{nw} = 0.5\ \text{nV}/\sqrt{\text{Hz}}$ and $f_{ce} = 1.5$ Hz.

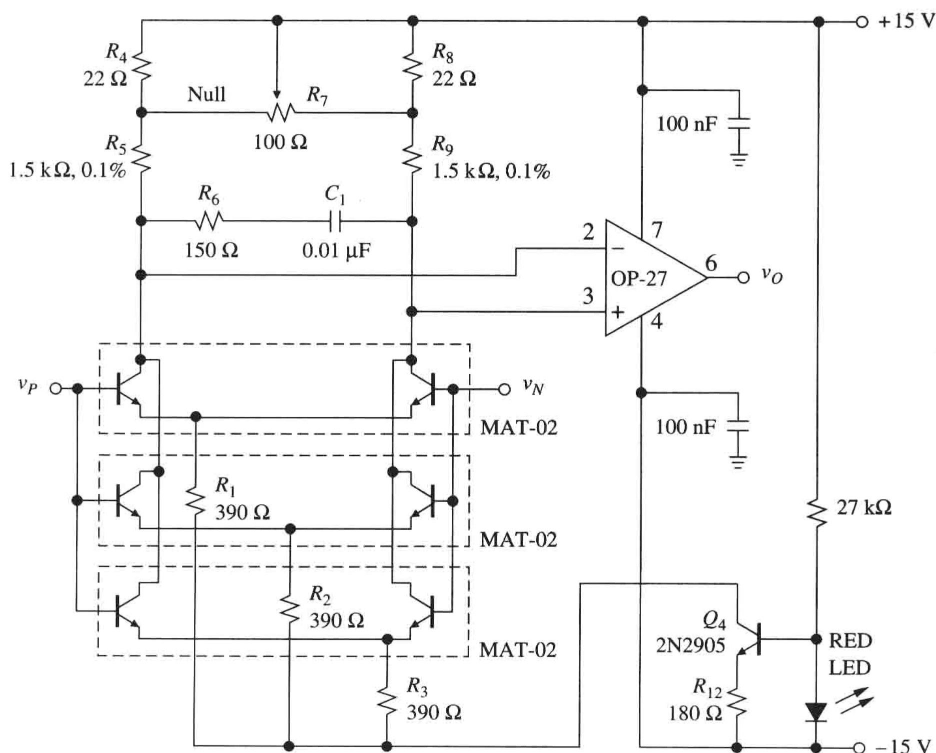


FIGURE 7.21
Ultralow-noise op amp. (Courtesy of Analog Devices.)

Transistor Q_4, in conjunction with R_{12} and the LED, forms a temperature-stable 6-mA current sink that R_1 through R_3 then split evenly among the three differential pairs. R_6 and C_1 provide frequency compensation for closed-loop gains greater than 10, and R_7 nulls the input offset voltage.

The additional gain provided by the front end increases the overall dc gain to $a_0 = 3 \times 10^7$ V/V. Other measured parameters are $i_{nw} = 1.5$ pA/$\sqrt{\text{Hz}}$, TC(V_{OS}) = 0.1 μV/°C (max), GBP = 150 MHz with $A_0 = 10^3$ V/V, and CMRR$_{\text{dB}}$ = 130 dB. Similar front-end designs can be used to improve the noise characteristics of other critical circuits, such as instrumentation amplifiers and audio preamps.

PROBLEMS

7.1 Noise properties

7.1 Two IC noise spot measurements, performed respectively at $f_1 = 10$ Hz and $f_2 \gg f_{ce}$, yield $e_n(f_1) = 20$ nV/$\sqrt{\text{Hz}}$ and $e_n(f_2) = 6$ nV/$\sqrt{\text{Hz}}$. Find the rms noise from 1 mHz to 1 MHz.

7.2 Consider two integrated-circuit noise sources e_{n1} and e_{n2} of the type of Eq. (7.8) in series with each other. (*a*) Show that the equivalent noise e_n of the combination of the two sources is still of the type of Eq. (7.8) by deriving expressions for its white-noise floor e_{nw} and corner frequency f_{ce} in terms of the individual noise floors and corner frequencies. (*b*) If e_{n1} has $e_{nw1} = 30$ nV/$\sqrt{\text{Hz}}$ and $f_{ce1} = 400$ Hz, and e_{n2} has $e_{nw2} = 40$ nV/$\sqrt{\text{Hz}}$ and $f_{ce2} = 100$ Hz, what is the total noise E_n from 1 Hz to 1 MHz?

7.2 Noise dynamics

7.3 (*a*) Using the integral tables (you can find them on the web), prove Eq. (7.11). (*b*) Repeat, but for Eq. (7.14).

7.4 (*a*) Assuming $f_L = 0$, find the NEB of the composite amplifier of Fig. 6.6*a*, consisting of two identical stages in cascade. How does this NEB compare with that of each individual stage? Justify in terms of Bode-plot areas. (*b*) How would you approximate for the flicker-noise frequency f_H?

7.5 (*a*) Consider the case $Q = 1/\sqrt{2}$ in Eq. (7.13*b*). Compare NEB$_{\text{BP}}$ with NEB$_{\text{LP}}$, and compare NEB$_{\text{LP}}$ with the NEB of a first-order circuit, and justify the differences intuitively in terms Bode-plot areas. (*b*) Compare NEB$_{\text{BP}}$ with NEB$_{\text{LP}}$ for the case $Q = 10$, and justify the differences.

7.6 Confirm the value of f_H of Example 7.2 via hand calculations. *Hint:* Apply Eq. (7.12), along with Eqs. (7.11*b*) and (7.13*b*).

7.7 (*a*) Find the NEB of a filter consisting of an *R-C* network, followed by a buffer, followed by another *R-C* network. (*b*) Repeat, but for a filter consisting of a *C-R* network, followed by a buffer, followed by an *R-C* network. (*c*) Repeat, but for a filter consisting of an *R-C* network, followed by a buffer, followed by a *C-R* network. (*d*) Rank the three filters in terms of noise minimization.

7.8 Estimate the NEB of the RIAA response of Fig. 3.13. Confirm with PSpice.

7.9 Find the NEB if $A_n(s)$ has two zeros at $s = -20\pi$ rad/s and $s = -2\pi\, 10^3$ rad/s, and four poles at $s = -200\pi$ rad/s, $s = -400\pi$ rad/s, $s = -2\pi\, 10^4$ rad/s, and $s = -2\pi\, 10^4$ rad/s.

7.10 Find the total output noise when a noise source with $f_{ce} = 100$ Hz and $e_{nw} = 10\,\text{nV}/\sqrt{\text{Hz}}$ is played through a noiseless wideband band-pass filter with a mid-frequency gain of 40 dB and −3-dB frequencies of 10 Hz and 1 kHz. Confirm via the pink-noise tangent.

7.11 A certain voltage reference is buffered to the outside via an FET-input op amp voltage follower. The reference's output noise e_{n1} has $e_{nw1} = 100\,\text{nV}/\sqrt{\text{Hz}}$ and $f_{ce1} = 20$ Hz, and the noise produced by the op amp itself is modeled by a source e_{n2} (in series with e_{n1}) and such that e_{n2} has $e_{nw2} = 25\,\text{nV}/\sqrt{\text{Hz}}$ and $f_{ce2} = 200$ Hz. Moreover, the op amp is dominant-pole compensated for GBP = 1 MHz via an external capacitance C_c. (*a*) Calculate the total noise above 0.1 Hz at the output of the follower. (*b*) Repeat if C_c is increased such that GBP = 10 kHz.

7.12 The spectral noise e_{no} of a certain amplifier below 100 Hz consists of $1/f$ noise with $f_{ce} = 1$ Hz and $e_{nw} = 10\,\text{nV}/\sqrt{\text{Hz}}$; from 100 Hz to 1 kHz it rolls off at the rate of −1 dec/dec; from 1 kHz to 10 kHz it is again constant at 1 nV/$\sqrt{\text{Hz}}$; and past 10 kHz it rolls off at the rate of −1 dec/dec. Sketch and label e_{no}, estimate the total rms noise above 0.01 Hz, and confirm using the pink-noise tangent principle.

7.3 Sources of noise

7.13 The noise of an IC current source is measured at two different frequencies and is found to be $i_n(250\text{ Hz}) = 6.71\,\text{pA}/\sqrt{\text{Hz}}$ and $i_n(2500\text{ Hz}) = 3.55\,\text{pA}/\sqrt{\text{Hz}}$. (*a*) What are the values of i_{nw} and f_{ci}? (*b*) If the source is fed to a 1-kΩ resistance, find the values of e_{nw} and f_{ce} for the noise voltage across its terminals. (*c*) Find the total rms noise E_n above 0.01 Hz if a 10-nF capacitor is placed across the resistor.

7.14 A diode with $I_s = 2$ fA, $r_S \cong 0$, $K = 10^{-16}$ A, and $a = 1$ is forward-biased at $I_D = 100\,\mu$A by a 3.3-V supply via a suitable series resistance R. The supply itself exhibits noise that can be assumed white with a density of $e_{ns} = 100\,\text{nV}/\sqrt{\text{Hz}}$. Combine the supply noise e_{ns}, the resistor noise e_{nr}, and the diode noise i_n, to find the overall noise voltage $e_n(f)$ across the diode, and express it in the form of Eq. (7.8). What are the values of e_{nw} and f_{ce}? *Hint:* To perform noise analysis, replace the diode with its ac equivalent, consisting of a noiseless resistance $r_d = V_T/I_D = (26\text{ mV})/I_D$.

7.15 The LT1009 2.5-V reference diode, when suitably biased, acts as a 2.5-V source with superimposed noise of the type $e_n^2 \cong (118\,\text{nV}/\sqrt{\text{Hz}})^2(30/f + 1)$. If the diode voltage is sent through an *R-C* filter with $R = 10$ kΩ and $C = 1\,\mu$F, estimate the peak-to-peak noise that one would observe at the output over a 1-minute interval.

7.16 Find a resistance that will produce the same amount of room-temperature noise as a diode operating with (*a*) a forward-bias current of 50 μA, and (*b*) a reverse-bias current of 1 pA.

7.17 (*a*) Show that the total rms noise voltage across the parallel combination of a resistance R and capacitance C is $E_n = \sqrt{kT/C}$, regardless of R. (*b*) Find an expression for the total rms value of the noise current flowing through a resistance R in series with an inductance L.

7.18 (*a*) Find a resistance that produces the same e_{nw} as a 741 op amp at room temperature. (*b*) Find a reverse-biased diode current that produces the same i_{nw} as a 741 op amp. How does this current compare with the input bias current of the 741?

7.4 Op amp noise

7.19 In the difference amplifier of Fig. 1.18 let $R_1 = R_3 = 10$ kΩ and $R_2 = R_4 = 100$ kΩ. Find the total output noise E_{no} above 0.1 Hz if the op amp is (*a*) the 741 type, and (*b*) the OP27 type. Compare also the individual components E_{noe}, E_{noi}, and E_{noR}, and comment. For the 741 assume $f_t = 1$ MHz, $e_{nw} = 20$ nV/$\sqrt{\text{Hz}}$, $f_{ce} = 200$ Hz, $i_{nw} = 0.5$ pA/$\sqrt{\text{Hz}}$, and $f_{ci} = 2$ kHz; for the OP27 assume $f_t = 8$ MHz, $e_{nw} = 3$ nV/$\sqrt{\text{Hz}}$, $f_{ce} = 2.7$ Hz, $i_{nw} = 0.4$ pA/$\sqrt{\text{Hz}}$, and $f_{ci} = 140$ Hz.

7.20 Using a 741 op amp, design a circuit that accepts three inputs v_1, v_2, and v_3, and yields $v_O = 2(v_1 - v_2 - v_3)$; hence, estimate its total output noise above 1 Hz.

7.21 In the bridge amplifier of Fig. P1.81 let $R = 100$ kΩ and $A = 2$ V/V, and let the op amps be 741 types. Estimate the total output noise above 1 Hz.

7.22 (*a*) Find the total rms output noise above 0.1 Hz for the *I-V* converter of Fig. 2.2 if $R = 10$ kΩ, $R_1 = 2$ kΩ, $R_2 = 18$ kΩ, and the op amp is the OP27, whose characteristics are given in Problem 7.19. (*b*) Find the SNR if i_I is a triangular wave with peak values of ± 10 μA.

7.23 (*a*) Find the total output noise above 0.1 Hz for the inverting amplifier of Fig. P1.64 if all resistances are 10 kΩ and the op amp is the 741 type. (*b*) Find the SNR if $v_I = 0.5 \cos 100t + 0.25 \cos 300t$ V.

7.24 A JFET-input op amp with $e_{nw} = 18$ nV/$\sqrt{\text{Hz}}$, $f_{ce} = 200$ Hz, and $f_t = 3$ MHz is configured as an inverting integrator with $R = 159$ kΩ and $C = 1$ nF. Estimate the total output noise above 1 Hz.

7.25 It is required to design an amplifier with $A_0 = 60$ dB using op amps with GBP $=$ 1 MHz. Two alternatives are being evaluated, namely, a single-op-amp realization and a two-op-amp cascade realization of the type of Example 6.2. Assuming the resistances are sufficiently low to render current and resistor noise negligible, which of the two configurations is noisier and by how much?

7.26 Using the OP227 dual op amp, design a dual-op amp instrumentation amplifier with a gain of 10^3 V/V, and find its total output noise above 0.1 Hz. Try keeping noise as low as practical. The OP227 consists of two OP27 op amps in the same package, so use the data of Problem 7.19.

7.27 With reference to the triple-op amp instrumentation amplifier of Fig. 2.20, consider the first stage, whose outputs are v_{O1} and v_{O2}. (*a*) Show that if OA_1 and OA_2 are dual op amps with densities e_n and i_n, the overall input power density of this stage is $e_{ni}^2 = 2e_n^2 + [(R_G \parallel 2R_3)i_n]^2/2 + 4kT(R_G \parallel 2R_3)$. (*b*) Estimate the total rms noise produced by this stage above 0.1 Hz if $R_G = 100$ Ω, $R_3 = 50$ kΩ, and the op amps are from the OP227 dual-op amp package, whose characteristics are the same as those of the OP27 given in Problem 7.19.

7.28 (*a*) In the triple-op amp instrumentation amplifier of Fig. 2.21 let the pot be adjusted for a gain of 10^3 V/V. Using the results of Problem 7.27, estimate the total output noise above 0.1 Hz. (*b*) Find the SNR for a sinusoidal input having a peak amplitude of 10 mV.

7.29 Use PSpice to verify the CFA noise calculations of Example 7.10.

7.30 The circuit of Fig. 7.9*a* has $R_1 = R_3 = 10\ \Omega$ and $R_2 = 10\ \text{k}\Omega$, and its output is observed through a band-pass filter having NEB = 100 Hz. The reading is 0.120 mV (rms), and it can be regarded as being primarily voltage noise since R_1 and R_3 are so small. Next, a 500-kΩ resistor is inserted in series with each input pin of the op amp to generate substantial current noise. The output reading is now 2.25 mV rms. Find e_n and i_n.

7.31 (*a*) Derive the transfer function of the noise filter of Fig. 7.13. (*b*) Modify the circuit so that it works as an inverting voltage amplifier with $H = -10H_{LP}$.

7.32 Using two 0.1-μF capacitances, specify resistances in the noise filter of Fig. 7.13 for $f_0 = 100$ Hz and $Q = 1/2$. If the op amp is the 741 type, find the total rms noise generated by the filter above 0.01 Hz with V_i and I_i both set to zero.

7.33 Using the voltage-input option of the noise filter of Fig. 7.13, design a circuit to filter the voltage of the LT1009 reference diode of Problem 7.15 for a total output noise above 0.01 Hz of 1 μV (rms) or less. Assume an OP27 op amp whose characteristics are given in Problem 7.19.

7.34 A student feeds the output of the LT1009 2.5-V reference diode (see Problem 7.15) to a voltage follower as well as to a unity-gain inverting amplifier so as to create a buffered dual reference of ±2.5 V. The op amps used are JFET types with GBP = 1 MHz, $e_{nw} = 10\ \text{nV}/\sqrt{\text{Hz}}$, and $f_{ce} = 250$ Hz, and the inverting amplifier uses two 20-kΩ resistors. Find the total output noise above 1 Hz for both the +2.5-V and the −2.5-V op amp outputs. Which one is less noisy, and why? Is this a good design, or can you suggest improvements?

7.5 Noise in photodiode amplifiers

7.35 Use PSpice to plot e_{noe}, e_{noi}, e_{noR}, and e_{no} for the circuit of Example 7.11. Hence, use the "`s`" and "`sqrt`" Probe functions to find E_{no}.

7.36 Investigate the effect of connecting an additional capacitance $C_f = 2$ pF in parallel with R_2 in the photodiode amplifier of Example 7.11. How does it affect noise? The signal-gain bandwidth?

7.37 Use PSpice to confirm Example 7.12.

7.38 Derive Eqs. (7.33) and (7.34).

7.39 Rework Example 7.11, but with R_2 replaced by a *T*-network with $R_2 = 1\ \text{M}\Omega$, $R_3 = 2\ \text{k}\Omega$, and $R_4 = 18\ \text{k}\Omega$, everything else staying the same. Comment on your findings.

7.40 Verify Example 7.13 via PSpice.

7.41 Modify the circuit of Example 7.13 to filter noise without significantly reducing the signal bandwidth. What is the total output noise of your circuit?

7.6 Low-noise op amps

7.42 A popular noise reduction technique is to combine N identical voltage sources in the manner of Fig. P7.42. (a) Show that if the noise of the resistors is negligible, the output density e_{no} is related to the individual source densities e_n as $e_{no} = e_n/\sqrt{N}$. (b) Find the maximum value of the resistances in terms of e_n so that the rms noise contributed by the resistances is less than 10% of the rms noise due to the sources.

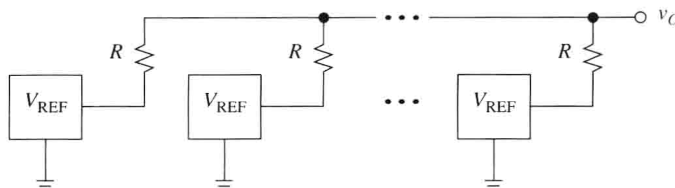


FIGURE P7.42

7.43 The circuit of Fig. P7.42 uses four LT1009 reference diodes (see Problem 7.15) and four 10-kΩ resistors, and includes also a 1-μF capacitor between the output node and ground. (a) Find the total output noise above 1 Hz. (b) Repeat, if three of the LT1009s are disconnected, leaving only one in place.

REFERENCES

1. H. W. Ott, *Noise Reduction Techniques in Electronic Systems,* 2d ed., John Wiley & Sons, New York, 1988.
2. C. D. Motchenbacher and J. A. Connelly, *Low-Noise Electronic System Design,* John Wiley & Sons, New York, 1993.
3. A. P. Brokaw, "An IC Amplifiers User's Guide to Decoupling, Grounding, and Making Things Go Right for a Change," Application Note AN-202, *Applications Reference Manual,* Analog Devices, Norwood, MA, 1993.
4. A. Rich, "Understanding Interference-Type Noise," Application Note AN-346, and "Shielding and Guarding," Application Note AN-347, *Applications Reference Manual,* Analog Devices, Norwood, MA, 1993.
5. F. N. Trofimenkoff, D. F. Treleaven, and L. T. Bruton, "Noise Performance of *RC*-Active Quadratic Filter Sections," *IEEE Trans. Circuit Theory*, Vol. CT-20, No. 5, September 1973, pp. 524–532.
6. A. Ryan and T. Scranton, "Dc Amplifier Noise Revisited," *Analog Dialogue,* Vol. 18, No. 1, Analog Devices, Norwood, MA, 1984.
7. M. E. Gruchalla, "Measure Wide-Band White Noise Using a Standard Oscilloscope," *EDN,* June 5, 1980, pp. 157–160.
8. P. R. Gray, P. J. Hurst, S. H. Lewis, and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 5th ed., John Wiley & Sons, New York, 2009, ISBN 978-0-470-24599-6.

9. S. Franco, "Current-Feedback Amplifiers," *Analog Circuits: World Class Designs*, R. A. Pease ed., Elsevier/Newnes, New York, 2008, ISBN 978-0-7506-8627-3.
10. W. Jung, *Op Amp Applications Handbook* (Analog Devices Series), Elsevier/Newnes, New York, 2005, ISBN 0-7506-7844-5.
11. R. M. Stitt, "Circuit Reduces Noise from Multiple Voltage Sources," *Electronic Design,* Nov. 10, 1988, pp. 133–137.
12. J. G. Graeme, *Photodiode Amplifiers–Op Amp Solutions,* McGraw-Hill, New York, 1996.
13. G. Erdi, "Amplifier Techniques for Combining Low Noise, Precision, and High Speed Performance," *IEEE J. Solid-State Circuits,* Vol. SC-16, December 1981, pp. 653–661.
14. A. Jenkins and D. Bowers, "NPN Pairs Yield Ultralow-Noise Op Amp," *EDN,* May 3, 1984, pp. 323–324.

8

STABILITY

自从 Harold S. Black 在 1927 年提出负反馈的概念以来，负反馈已经成为电子科学和控制科学，以及其他领域的应用科学（例如生物系统建模）的基础。正如前面几章提到的，负反馈能够改善多种性能，包括抑制生产和环境变化导致的增益不稳定性，减少由于器件非线性产生的失真，频带扩展和阻抗变换。负反馈的优点在高增益的放大器（例如运算放大器）中更加明显。

然而，负反馈的引入也有一定的代价，可能会导致振荡。一般来说，如果无论输入端输入什么，系统环路中都能够产生一个稳定的信号，那么就说产生了振荡。为了产生振荡，系统必须在环路中产生足够的相位偏移，使得负反馈变成正反馈。同时还要产生足够的环路增益，使得不加任何输入的情况下，还能产生输出振荡。

本章重点

本章系统地分析了导致不稳定的原因，并给出了适当的解决办法来使电路

稳定，也就是所谓的频率补偿技术。这使得负反馈的优点得到了充分利用。本章的开始给出了许多与稳定性相关的概念，比如相位裕度、增益裕度、峰值和振铃、截止速率、借助返回比分析环路增益以及借助电压 / 电流节点进行环路增益测量。

接下来，举例说明一些最常用的内部频率补偿的方案：主极点补偿、零极点补偿、密勒补偿、右半平面零点控制补偿和前馈补偿。本章还将进一步介绍含有反馈极点的电路，特别是杂散输入电容和容性负载的影响。接下来，本章还将讨论一些流行的补偿方案，比如输入滞后和反馈超前补偿。在研究了电流反馈放大器的稳定性之后，本章将以复合放大器结尾。

稳定性在计算机仿真领域得到了广泛的应用，不仅引入了一个新的指标，而且可以当做手动计算估计的验证工具。本章还广泛地应用 PSpice 来对之前的补偿方案进行直观验证。

8.1
THE STABILITY PROBLEM

The advantages of negative feedback are realized only if the circuit has been stabilized against the possibility of oscillations. For an intuitive discussion,[1,2] refer to the feedback system of Fig. 8.1, where for simplicity we assume unilateral blocks and no feedthrough, so we can apply Eq. (6.13),

$$A(jf) = \frac{V_o}{V_i} = A_{\text{ideal}} \times D(jf) \tag{8.1}$$

where A_{ideal} is the closed-loop gain in the limit $a \to \infty$, and

$$D(jf) = \frac{1}{1 + 1/T(jf)} \tag{8.2}$$

is the *discrepancy function*, and

$$T(jf) = a(jf)\beta(jf) \tag{8.3}$$

is the *loop gain*, which we find via the return-ratio techniques of Section 1.7. As we know, whenever the amplifier detects an input error V_d, it tries to reduce it. It takes

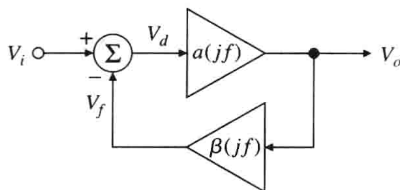


FIGURE 8.1
Negative-feedback system with unilateral amplifier and unilateral feedback network.

some time, however, for the amplifier to react and then transmit its response back to the input via the feedback network. The consequence of this combined delay is a tendency on the part of the amplifier to overcorrect the input error, especially if the loop gain is high. If the overcorrection exceeds the original error, a regenerative effect results, whereby the magnitude of V_d diverges, instead of converging, and instability results. Signal amplitudes grow exponentially until inherent circuit nonlinearities limit further growth, forcing the system either to saturate or to oscillate, depending on the order of its system function. By contrast, a circuit that succeeds in making V_d converge is stable.

Gain Margin

Whether a system is stable or unstable is determined by the manner in which its loop gain T varies with frequency. To substantiate, suppose a frequency exists at which the phase angle of T is $-180°$; call this frequency $f_{-180°}$. Then, $T(jf_{-180°})$ is real and negative, such as -0.5, -1, -2, indicating that feedback has turned from negative to positive. We consider three significant cases:

If $|T(jf_{-180°})| < 1$, then Eq. (1.40), rewritten as

$$A(jf_{-180°}) = \frac{a(jf_{-180°})}{1 + T(jf_{-180°})}$$

indicates that $A(jf_{-180°})$ is greater than $a(jf_{-180°})$ because the denominator is less than unity. The circuit is nonetheless stable because any signal circulating around the loop will progressively decrease in magnitude and eventually die out; consequently, the poles of $A(s)$ must lie in the left half of the s plane.

If $T(jf_{-180°}) = -1$, the above equation predicts $A(jf_{-180°}) \to \infty$, indicating that the circuit can now sustain an output signal even with zero input! The circuit is an oscillator, indicating that $A(s)$ must have a conjugate pole pair right on the imaginary axis. Oscillations are initiated by ac noise, which is always present in some form at the amplifier input. An ac noise component V_d right at $f = f_{-180°}$ results in a feedback component $V_f = -V_d$, which is further multiplied by -1 in the summing network to yield V_d itself. Thus, once this ac component has entered the loop, it will be sustained indefinitely.

If $|T(jf_{-180°})| > 1$, mathematical tools other than the foregoing equation are needed to predict circuit behavior. Suffice it to say here that now $A(s)$ may have a conjugate pole pair in the right half of the s plane. Consequently, once started, oscillation will grow in magnitude until some circuit nonlinearity, either inherent, such as a nonlinear VTC, or deliberate, such as an external clamping network, reduces the loop gain to exactly unity. Henceforth, oscillation is of the sustained type.

A quantitative measure of stability is offered by the *gain margin,* defined as

$$\mathrm{GM} = 20 \log \frac{1}{|T(jf_{-180°})|} \tag{8.4}$$

The GM represents the number of decibels by which we can increase $|T(jf_{-180°})|$ before it becomes unity and thus leads to instability. For instance, a circuit with $|T(jf_{-180°})| = 1/\sqrt{10}$ has $\mathrm{GM} = 20 \times \log_{10} \sqrt{10} = 10$ dB, which is considered a reasonable margin. By contrast, a circuit with $|T(jf_{-180°})| = 1/\sqrt{2}$ has GM $= 3$ dB,

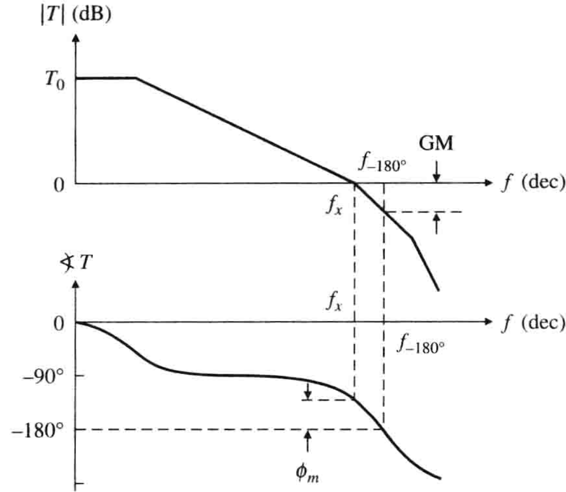


FIGURE 8.2
Visualizing gain margin GM and phase margin ϕ_m.

not much of a margin: only a modest increase in the gain a because of manufacturing process variations or environmental changes may easily lead to instability! The GM is visualized in Fig. 8.2 (top).

Phase Margin

An alternative and more common way of quantifying stability is via phase. In this case we focus on $\measuredangle T(jf_x)$, the phase angle of T at the crossover frequency f_x, where $|T| = 1$ by definition, and we define the *phase margin* ϕ_m as the number of degrees by which we can lower $\measuredangle T(jf_x)$ before it reaches $-180°$ and thus leads to instability. We have $\phi_m = \measuredangle T(jf_x) - (-180°)$, or

$$\phi_m = 180° + \measuredangle T(jf_x) \tag{8.5}$$

The phase margin is visualized in Fig. 8.2 (bottom). As we move along, we shall be interested in the quantity $|D(jf_x)|$. Since $\measuredangle T(jf_x) = \phi_m - 180°$, we have $T(jf_x) = 1\exp[j(\phi_m - 180°)] = -1\exp(j\phi_m)$. Substituting into Eq. (8.2) gives

$$|D(jf_x)| = \left|\frac{1}{1 + 1/(-1e^{j\phi_m})}\right| = \left|\frac{1}{1 - e^{-j\phi_m}}\right| = \left|\frac{1}{1 - (\cos\phi_m - j\sin\phi_m)}\right|$$
$$= \frac{1}{\sqrt{(1 - \cos\phi_m)^2 + (\sin\phi_m)^2}}$$

where Euler's formula has been used. Expanding and using the identity $\cos^2\phi_m + \sin^2\phi_m = 1$, we get

$$|D(jf_x)| = \frac{1}{\sqrt{2(1 - \cos\phi_m)}} \tag{8.6}$$

LAPLACE {v(I, F)} = {100k/((1+s/6.283E3)*(1+s/6.283E5)*(1+s/6.283E7))}

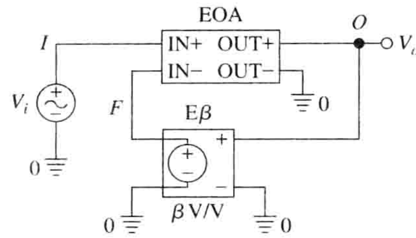


FIGURE 8.3
PSpice circuit to investigate a three-pole op amp under different amounts of feedback.

An Illustrative Example

Let us illustrate the above concepts using the feedback circuit of Fig. 8.3 as a vehicle. The circuit is based on a three-pole op amp with the open-loop gain

$$a(jf) = \frac{10^5}{(1 + jf/10^3)\,(1 + jf/10^5)\,(1 + jf/10^7)} \tag{8.7}$$

Magnitude and phase are calculated as

$$|a(jf)| = \frac{10^5}{\sqrt{\left[1 + (f/10^3)^2\right] \times \left[1 + (f/10^5)^2\right] \times \left[1 + (f/10^7)^2\right]}} \tag{8.8a}$$

$$\measuredangle a(jf) = -\left[\tan^{-1}(f/10^3) + \tan^{-1}(f/10^5) + \tan^{-1}(f/10^7)\right] \tag{8.8b}$$

and are plotted via PSpice in Fig. 8.4*a* (magnitude is in dB, as marked at the left, and phase in degrees, as marked at the right).

If the roots are widely spaced apart as in the present example, we can combine magnitude and phase in the more concise and visually intuitive form of Fig. 8.4*b*.

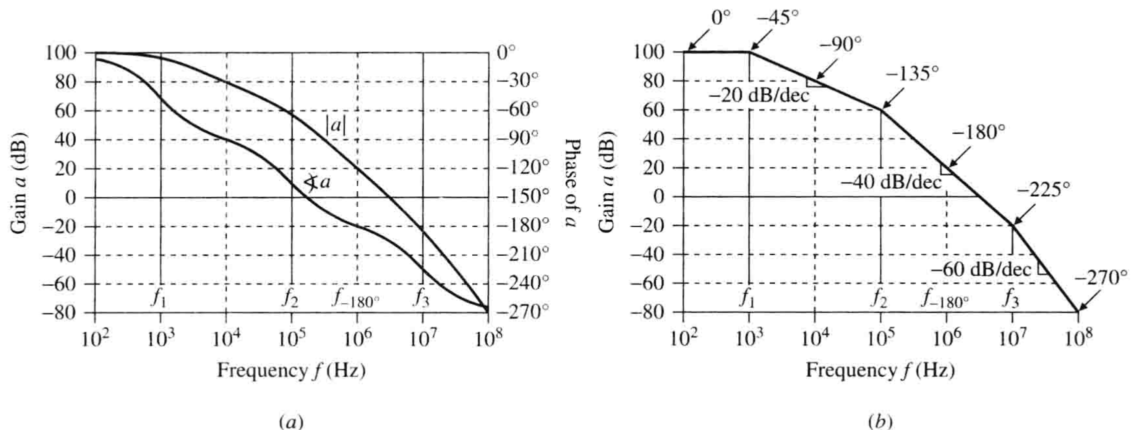


FIGURE 8.4
(*a*) Open-loop magnitude and phase plots for the op amp of Fig. 8.3 (*b*) Linearized magnitude plot associating phase with slope.

Specifically, we draw a linearized magnitude plot using segments of progressively steeper slopes, and we mark significant phase values using the correspondence

$$\text{Phase (in degrees)} \leftrightarrow 4.5 \times \text{Slope (in dB/dec)} \tag{8.9}$$

Thus, from dc to f_1 we draw a segment with a slope of 0 dB/dec, for which Eq. (8.9) implies a phase of 0°. From f_1 to f_2 we draw a segment with a slope of −20 dB/dec, implying a phase of $4.5 \times (-20)$, or −90°. Right at f_1 the slope is the mean of 0 dB/dec and −20 dB/dec, or −10 dB/dec, so the corresponding phase is $4.5 \times (-10)$, or −45°. Likewise, the segment from f_2 to f_3 has a slope of −40 dB/dec, implying a phase of −180°. The phase at f_2 is −135°, and past f_3 slope approaches −60 dB/dec and phase approaches −270°.

We now wish to investigate the closed-loop response under increasing amounts of frequency-independent feedback, in which case $\measuredangle T(jf) = \measuredangle a(jf)$. We have the following significant cases:

- Starting out with $\beta = 10^{-4}$, we observe that if we draw the $1/\beta$ line ($1/\beta = 10^4 =$ 80 dB) in Fig. 8.4*b*, it will intersect the gain curve at $f_x \cong 10$ kHz, where $\measuredangle a(jf_x) \cong -90°$. Consequently, $\phi_m \cong 180° - \measuredangle T(jf_x) = 180° - \measuredangle a(jf_x) = 180° - 90° = 90°$. The closed-loop response, shown in Fig. 8.5, exhibits a dominant pole frequency of f_x. The response to an input step of $\beta V (= 0.1 \text{ mV})$, shown in Fig. 8.6*a*, is an approximately *exponential transient*. Once the input is stepped back to zero, $v_O(t)$ will decay in approximately exponential fashion, indicating a stable circuit.
- Raising β from 10^{-4} to 10^{-3} would lower the $1/\beta$ line to 60 dB in Fig. 8.4*b*, giving $f_x \cong 100$ kHz and $\measuredangle a(jf_x) \cong -135°$, so $\phi_m \cong 180° - 135° = 45°$. Now the closed-loop gain exhibits a bit of *peaking* just before f_x, after which it rolls off with frequency as $|a(jf)|$. By Eq. (8.6) we have $|D(jf_x)| = 1/\sqrt{2(1-\cos 45°)} = 1.307$,

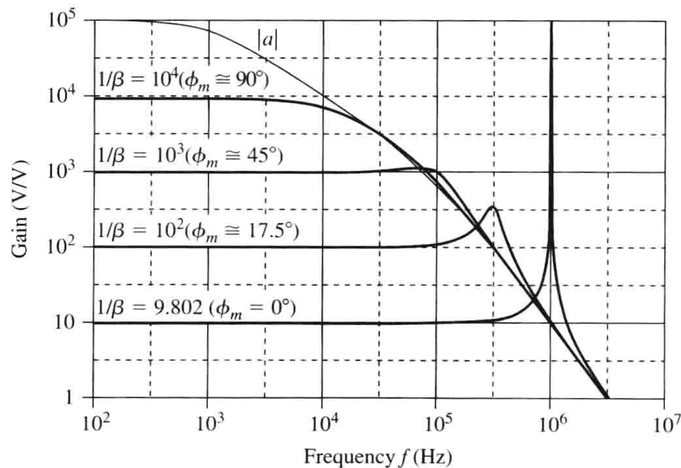


FIGURE 8.5
Closed-loop responses of the circuit of Fig. 8.3 for different amounts of feedback. Raising β lowers the $1/\beta$ curve, shifting the crossover frequency to regions of greater phase shift and thus lower phase margin. This, in turn, increases the amount of peaking as well as ringing (see Fig. 8.6).

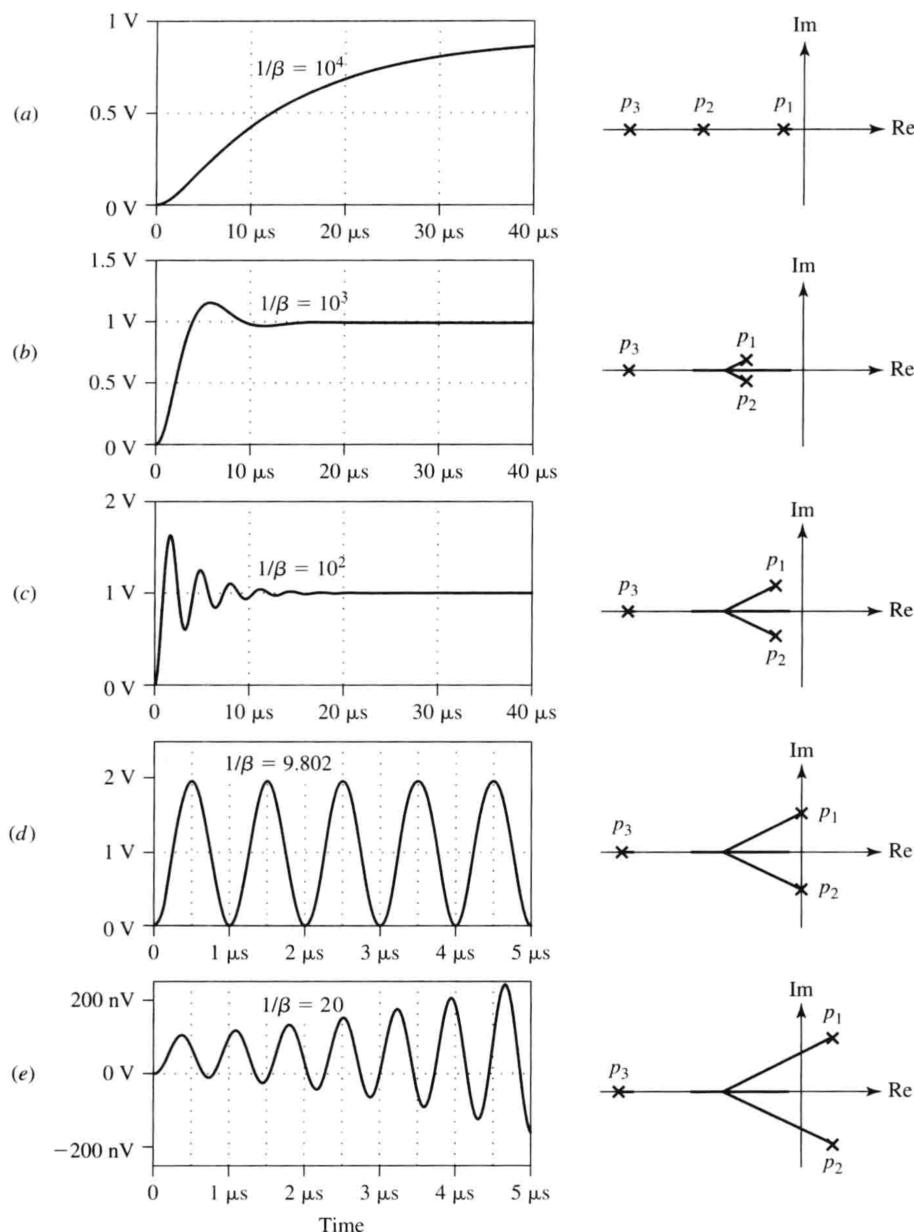


FIGURE 8.6
Step responses and s-plane pole locations of the circuit of Fig. 8.3 for increasing β.

so Eq. (8.1) predicts $|A(jf_x)| = 10^3 \times 1.307$, which is 30.7% *higher* than A_{ideal}. Recall from systems theory that *peaking* in the frequency domain is accompanied by *ringing* in the time domain. This is confirmed by Fig. 8.6*b*, showing the transient response to an input step of β V (= 1 mV). Once we step the input back to zero, $v_O(t)$ will decay to zero, albeit with a bit of ringing. We conclude that

a circuit with $\phi_m = 45°$ is still a stable circuit, though its (moderate) amount of peaking and ringing may be undesirable in certain applications.

- Raising β to 10^{-2} would lower the $1/\beta$ line to 40 dB in Fig. 8.4*b*, so f_x is now the *geometric mean* of 100 kHz and 1 MHz, or $f_x \cong \sqrt{10^5 \times 10^6} = 316$ kHz, and $\measuredangle a(jf_x)$ is the *arithmetic mean* of $-145°$ and $-180°$, or about $-162.5°$, so $\phi_m \cong 180° - 162.5° = 17.5°$. With a reduced phase margin, both peaking and ringing are more pronounced. In fact we now have $|D(jf_x)| = 1/\sqrt{2(1 - \cos 17.5°)} = 3.287$, so Eq. (8.1) predicts $|A(jf_x)| \cong 10^2 \times 3.28$, or almost $3.3 \times A_{ideal}$! Figure 8.6*c* shows the response to an input step of β V (= 10 mV). Once we step the input back to zero, $v_O(t)$ will decay with quite a bit of ringing. We conclude that a circuit with $\phi_m = 17.5°$, though still stable, is likely to be unacceptable in most applications because of excessive peaking and ringing.
- Cursor measurements on the PSpice plots of Fig. 8.4*a* give $f_{-180°} = 1.006$ MHz, where $|a(jf_{-180°})| = 9.802$ V/V, so if we let $\beta = 1/9.802 = 1.02 \times 10^{-1}$, we get $f_x = 1.006$ MHz and $\measuredangle a(jf_x) = -180°$, or $\phi_m = 0°$. Consequently, $|D(jf_x)|$ shoots up to infinity, indicating oscillatory behavior. This is confirmed by Fig. 8.6*d*, showing the response to a 100-mV input step.
- Raising β to 2×10^{-1} lowers the $1/\beta$ line further, pushing f_x into a region of additional phase shift and thus $\phi_m < 0°$. All it takes now is internal noise to trigger a growing oscillation. Using an input step of just 1 nV to simulate noise, we obtain the response of Fig. 8.6*e*.

It is instructive to visualize circuit behavior also in terms of its poles in the complex plane. Letting $jf \to (s/2\pi)$ in Eq. (8.7), substituting into Eq. (8.2) and then into Eq. (8.1), we get, after minor algebra,

$$A(s) = \frac{10^5}{\beta 10^5 + \left(1 + \dfrac{s}{2\pi 10^3}\right)\left(1 + \dfrac{s}{2\pi 10^5}\right)\left(1 + \dfrac{s}{2\pi 10^7}\right)}$$

The roots of the denominator are the poles of $A(s)$. Using a scientific calculator or similar, we find the poles of Table 8.1. These data are best visualized in the s-plane as shown (not to scale) at the right side of Fig. 8.6. Starting out with no feedback ($\beta = 0$) and gradually increasing β brings the two lowest poles closer together, until they become coincident and then split apart to become complex conjugate and move toward the imaginary axis. Once on the imaginary axis, they result in sustained

TABLE 8.1
Closed-loop poles for the circuit of Fig. 8.3

β	$p_1(s^{-1})$	$p_2(s^{-1})$	$p_3(s^{-1})$
0	$2\pi(-1.0\text{ k})$	$2\pi(-100\text{ k})$	$2\pi(-10\text{ M})$
10^{-4}	$2\pi(-12.4\text{ k})$	$2\pi(-88.5\text{ k})$	$2\pi(-10\text{ M})$
10^{-3}	$2\pi(-50\text{ k} + j87.2\text{ k})$	$2\pi(-50\text{ k} - j87.2\text{ k})$	$2\pi(-10\text{ M})$
10^{-2}	$2\pi(-45.4\text{ k} + j313\text{ k})$	$2\pi(-45.4\text{ k} - j313\text{ k})$	$2\pi(-10.01\text{ M})$
1.02×10^{-1}	$2\pi(0 + j1.0\text{ M})$	$2\pi(0 - j1.0\text{ M})$	$2\pi(-10.1\text{ M})$
2×10^{-1}	$2\pi(46.7\text{ k} + j1.34\text{ M})$	$2\pi(46.7\text{ k} - j1.34\text{ M})$	$2\pi(-10.2\text{ M})$

oscillation; once they spill into the right half of the complex plane, they result in a growing oscillation.

Looking at Fig. 8.5, we observe that if we can tolerate the amount of peaking that comes with, say, $\phi_m = 45°$, then we must restrict operation to $1/\beta \geq 10^3$ V/V. What if we want to operate the amplifier at lower gains, such as $1/\beta = 50$ V/V or $1/\beta = 2$ V/V? As is, the circuit will just oscillate with these βs! Mercifully, clever frequency-compensation techniques have been developed that allow us to stabilize an amplifier for virtually any gain we wish, including what by now we recognize as the most difficult configuration to stabilize, namely, the voltage follower, for which $\beta = 1$.

EXAMPLE 8.1. (*a*) What is the minimum allowable noise gain $1/\beta$ if we want to operate the amplifier of Fig. 8.3 with a phase margin of 60°? (*b*) Find $|D(jf_x)|$ and comment. (*c*) Verify with PSpice.

Solution.

(*a*) By Eq. (8.5) we have $\measuredangle T(jf_x) = \phi_m - 180° = 60° - 180° = -120°$. A look at Fig. 8.4*a* indicates that $f_{-120°}$ is about an octave below 100 kHz. Start out with the initial estimate $f_{-120°} = 50$ kHz and iterate using Eq. (8.8*b*) until you settle at $f_{-120°} = 59.2$ kHz. Next, use Eq. (8.8*a*) to find that $|a(jf_{-120°})| \cong 1,453$ V/V. This is the minimum value of $1/\beta$ for $\phi_m \geq 60°$ with this particular amplifier.

(*b*) Proceeding in the usual manner, we get

$$|D(jf_x)| \cong \frac{1}{\sqrt{2\,(1 - \cos 60°)}} = 1$$

Since at dc we have $T_0 = 10^5/1,453 = 68.8$, it follows that $D_0 = 1/(1+1/68.8) = 0.986 < |D(jf_x)|$, indicating a very small amount of peaking.

(*c*) Using the circuit of Fig. 8.3 with $\beta = 1/1,453 = 0.688 \times 10^{-3}$ V/V we get the plots of Fig. 8.7. It is fair to say that, aside from a slight amount of peaking and ringing, which is acceptable in many applications, the condition $\phi_m = 60°$ is *almost as good* as $\phi_m = 90°$ in terms of stability, yet it expands the range of possible closed-loop gains by decreasing the lower limit of acceptable values for $1/\beta$ from 10^4 V/V to 1,453 V/V, or by almost 17 dB.

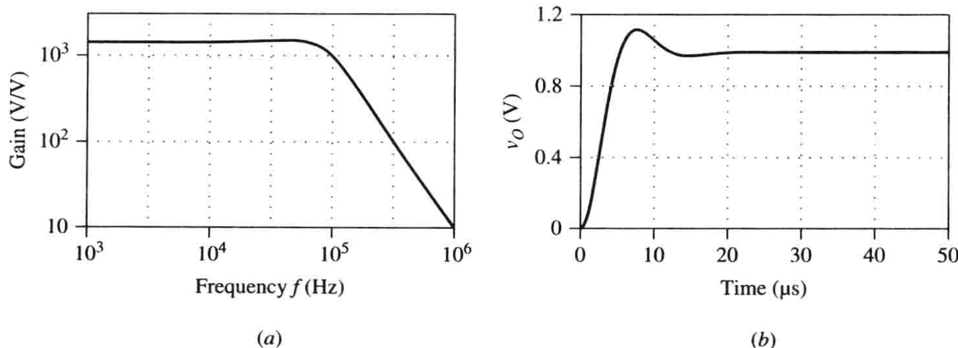


FIGURE 8.7
(*a*) Frequency and (*b*) step responses for the special case $\phi_m = 60°$.

Peaking and Ringing as Functions of the Phase Margin ϕ_m

Peaking in the frequency domain and ringing in the time domain are quantified, respectively, via the *gain peaking GP*, in dB, and the *overshoot OS, in percentage*, in the manner depicted in Fig. 8.8. Both effects are absent in first-order systems since it takes a complex pole pair to produce them. For a second-order all-pole system, peaking occurs for $Q > 1/\sqrt{2}$, and ringing for $\zeta < 1$, where the *quality factor Q* and the *damping ratio* ζ are related as $Q = 1/2\zeta$, or $\zeta = 1/(2Q)$. Second-order all-pole systems are well documented in the literature,[3] where it is found that

$$\text{GP} = 20 \log_{10} \frac{2Q^2}{\sqrt{4Q^2 - 1}} \qquad \text{for } Q > 1/\sqrt{2} \tag{8.10}$$

$$\text{OS}\,(\%) = 100 \exp \frac{-\pi\zeta}{\sqrt{1-\zeta^2}} \qquad \text{for } \zeta < 1 \tag{8.11}$$

$$\phi_m = \cos^{-1}\left(\sqrt{4\zeta^4 + 1} - 2\zeta^2\right) = \cos^{-1}\left(\sqrt{1 + \frac{1}{4Q^4}} - \frac{1}{2Q^2}\right) \tag{8.12}$$

Combining these equations yields the graphs of Fig. 8.9, which relate peaking and ringing to the phase margin. We observe that peaking occurs for $\phi_m \le \cos^{-1}(\sqrt{2} - 1) = 65.5°$, and ringing for $\phi_m \le \cos^{-1}(\sqrt{5} - 2) = 76.3°$. It is also worth keeping in mind the following frequently encountered values of GP (ϕ_m) and OS (ϕ_m):

$$\text{GP}\,(60°) \cong 0.3 \text{ dB} \qquad \text{OS}\,(60°) \cong 8.8\%$$

$$\text{GP}\,(45°) \cong 2.4 \text{ dB} \qquad \text{OS}\,(45°) \cong 23\%$$

Depending on the case, a closed-loop response may have a single pole, a pole pair, or a higher number of poles. Mercifully, the response of higher-order circuits is often dominated by a single pole pair, so the graphs of Fig. 8.9 provide a good starting point for a great many circuits of practical interest.

The Rate of Closure (ROC)

In everyday's practice an engineer must be able to quickly assess the degree of stability of a circuit. A popular tool for circuits with poles (and possibly zeros) in

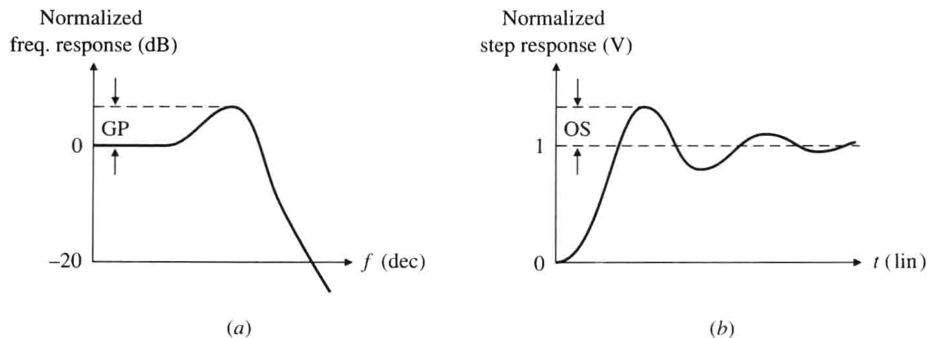


FIGURE 8.8
Illustrating gain peaking GP and overshoot OS.

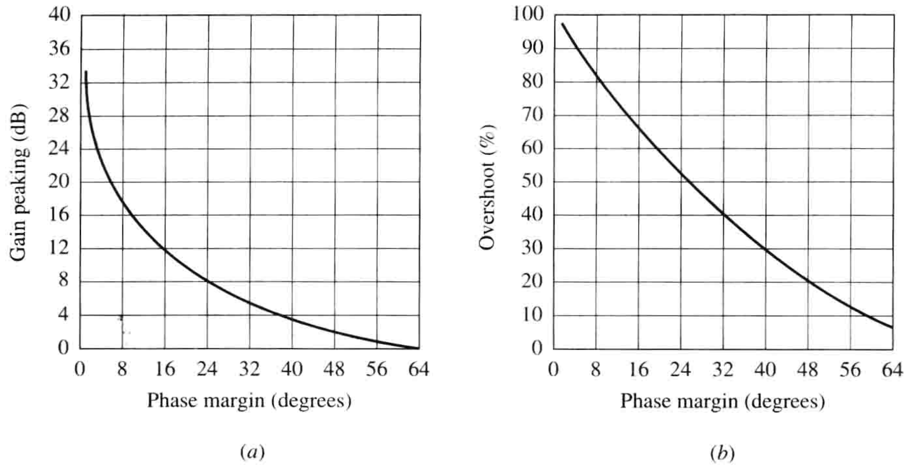


FIGURE 8.9
GP and OS as functions of ϕ_m for a second-order all-pole system.

the left-half of the complex plane (such circuits are referred to as *minimum-phase circuits*) is offered by the rate of closure (ROC), representing the difference between the slopes of the $|1/\beta|$ and $|a|$ curves right at the crossover frequency,

$$\text{ROC} = \text{Slope of } \left| \frac{1}{\beta(jf_x)} \right| - \text{Slope of } |a(jf_x)| \tag{8.13}$$

Once we know the ROC, we use Eq. (8.9) to estimate the phase margin as

$$\phi_m (\text{in degrees}) \cong 180^\circ - 4.5 \times \text{ROC (in dB/dec)} \tag{8.14}$$

Let us illustrate via the typical cases depicted in Fig. 8.10.

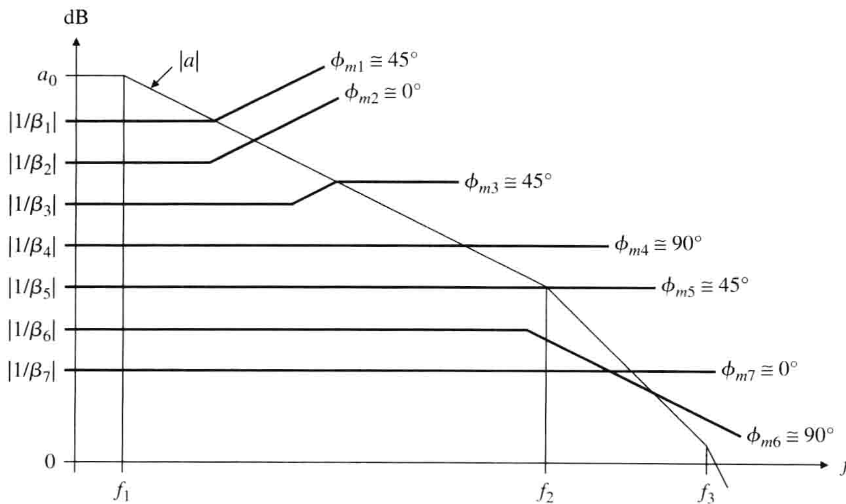


FIGURE 8.10
Rate of closure (ROC) for different feedback-factor types.

The $|1/\beta_4|$, $|1/\beta_5|$, and $|1/\beta_7|$ curves are flat, indicating frequency-independent feedback of the types already seen in connection with Fig. 8.5. Thus, for $|1/\beta_4|$, Eq. (8.13) gives $\text{ROC}_4 = 0 - (-20) = 20$ dB/dec, so $\phi_{m4} \cong 180 - 4.5 \times 20 = 90°$. Likewise, for $|1/\beta_5|$ we have $\text{ROC}_5 = 0 - (-30) = 30$ dB/dec, so $\phi_{m5} \cong 180 - 4.5 \times 30 = 45°$; for $|1/\beta_7|$ we have $\text{ROC}_7 = 0 - (-40) = 40$ dB/dec, so $\phi_{m7} \cong 180 - 4.5 \times 40 = 0°$.

The $|1/\beta_1|$ and $|1/\beta_2|$ curves refer to the case of a zero frequency for $1/\beta$, and thus a pole frequency for β. A pole inside the loop introduces *phase lag*, thus eroding the phase margin. In fact, the $|1/\beta_2|$ curve has $\text{ROC}_2 = +20 - (-20) = 40$ dB/dec, so $\phi_{m2} \cong 180 - 4.5 \times 40 = 0°$. The phase margin is higher for the $|1/\beta_1|$ curve, whose break point was made to coincide with the crossover frequency so that $\text{ROC}_1 = 10 - (-20) = 30$ dB/dec, thus resulting in $\phi_{m1} \cong 45°$.

The $|1/\beta_6|$ curve exhibits a pole frequency, or a zero frequency for β_6. A left-plane zero inside the loop introduces *phase lead*, thus ameliorating ϕ_m. In fact, $\text{ROC}_6 = -20 - (-40) = 20$ dB/dec, so $\phi_{m6} \cong 90°$. (In this circuit β_6 is more stable than β_7, even though $f_{x6} > f_{x7}$!). It pays to regard ROC as the *angle* between the $|1/\beta|$ and $|a|$ curves at f_x. The lower this angle, the more stable the circuit.

8.2 PHASE AND GAIN MARGIN MEASUREMENTS

In the course of designing a negative-feedback circuit, we need to monitor its phase margin ϕ_m to ensure that the stability specifications are met (if they aren't, we need to take suitable frequency-compensation measures). As an example, let us investigate the circuit of Fig. 8.11, consisting of an inverting amplifier driving a capacitive load R_L-C_L, and inclusive also of the inverting-input stray capacitance C_n. Each of C_L and C_n creates a pole within the loop, so we expect a two-fold erosion of the phase margin.

Return-Ratio Analysis

If the op amp lends itself to be modeled via its Thevenin or Norton equivalent, we use *return-ratio analysis* to plot $T(jf)$. Next, we find f_x, measure $\measuredangle T(jf_x)$, and finally let $\phi_m = 180° + \measuredangle T(jf_x)$. Alternatively, we find $f_{-180°}$, measure $|T(jf_{-180°})|$, and let $\text{GM} = -20 \log |T(jf_{-180°})|$.

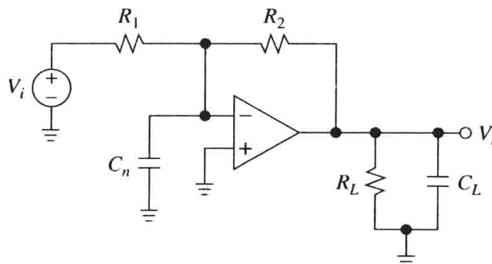


FIGURE 8.11
An illustrative circuit example for stability evaluation.

EXAMPLE 8.2. Let the op amp of Fig. 8.11 have $a_0 = 10^5$ V/V, $r_d = 1$ MΩ, $r_o = 100$ Ω, and two pole frequencies of 10 Hz and 2 MHz. (*a*) Use PSpice to find ϕ_m and GM if $R_2 = 2R_1 = 100$ kΩ, $R_L = 2$ kΩ, $C_n = 3$ pF, and $C_L = 1$ nF. (*b*) What is the phase-margin erosion due to C_n? That due to C_L? (*c*) What is ϕ_m if $C_n = C_L = 0$? Comment on your findings.

Solution.

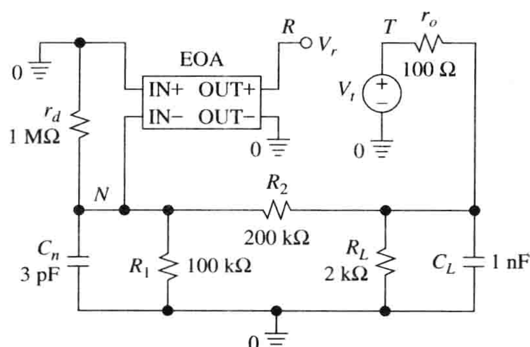
(*a*) As shown in Fig. 8.12*a*, we model the open-loop response with a dependent source of the Laplace type. Next, we set the input to zero, we break the loop right at the Laplace source's output to inject a test voltage V_t downstream of the source, we find the voltage V_r returned by the source, and we let $T(jf) = -V_r/V_t$. The resulting magnitude and phase plots are shown in Fig. 8.12*b*. Using PSpice's cursor facility, we find $f_x = 276.4$ kHz and $\measuredangle T(jf_x) = -125.3°$, so $\phi_m = 180 - 125.3 = 54.7°$. Also, $f_{-180°} = 792$ kHz and $|T(jf_{-180°})| = -12.7$ dB, so GM = 12.7 dB.

(*b*) Rerunning with $C_n = 0$, we get $f_x = 290.1$ kHz and $\measuredangle T(jf_x) = -108.1°$, so $\phi_m = 180 - 108.1 = 71.9°$. The presence of C_n erodes ϕ_m by $71.9 - 54.7 = 17.2°$. Rerunning with $C_n = 3$ pF and $C_L = 0$ gives $f_x = 279.8$ kHz and $\measuredangle T(jf_x) = -116.2°$, or $\phi_m = 180 - 116.2 = 63.8°$, so the phase erosion by C_L is $63.8 - 54.7 = 9.1°$.

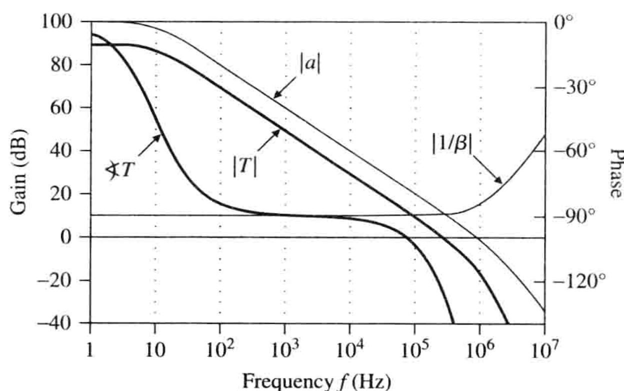
(*c*) Rerunning with $C_n = C_L = 0$, we get $f_x = 294.3$ kHz and $\measuredangle T(jf_x) = -98.4°$, so $\phi_m = 81.6°$. In this case the $1/|\beta|$ curve is flat and $\measuredangle T(jf_x)$ consists of $-90°$ due to the 10-Hz op amp pole, and $-8.4°$ due to the 2-MHz pole. Clearly, the presence of C_n and C_L results in a pole-frequency pair for β, or a zero-frequency pair for $1/\beta$, causing a high-frequency two-fold upward bend for the $1/|\beta|$ curve (see Fig. 8.12*b*). This increases the ROC, in turn reducing ϕ_m.

Double-Injection Techniques

Return-ratio analysis postulates a dependent-source model for the op amp. If the op amp is available in transistor-level or in macromodel-level form, we do not have access to such a source, so we must use other means to measure ϕ_m. An elegant



(*a*)



(*b*)

FIGURE 8.12
(*a*) PSpice circuit to find the loop gain $T(jf)$ via return-ratio analysis, and (*b*) Bode plots. The $|a|$ trace is plotted as `DB(V(R)/(-V(N)))`, the $|1/\beta|$ trace as `DB(V(T)/V(N))`, the $|T|$ trace as `DB(-V(R)/V(T))`, and the $\measuredangle T$ trace as `P(-V(R)/V(T))`.

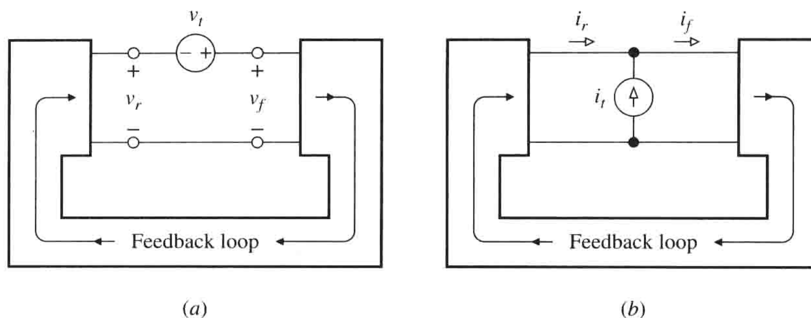


FIGURE 8.13
Illustrating (*a*) voltage injection and (*b*) current injection.

alternative for finding $T(jf)$, which is suited both to SPICE simulation and to testing in the lab, is the *successive voltage* and *current injection technique* devised by R. D. Middlebrook[4] and depicted in Fig. 8.13. The steps are as follows: First, set all external signal sources to zero so as to put the circuit in its *dormant* state. Next, break the feedback loop and insert a *series* test-source v_t as in Fig. 8.13*a*. The perturbation introduced by v_t causes a signal v_f to propagate in the *forward* direction, to which the loop responds with a *return* signal v_r. Let

$$T_v = -\frac{v_r}{v_f} \tag{8.15a}$$

Next, remove the test source v_t and apply, between the *same* set of wires, a *shunt* test source i_t as shown in Fig. 8.13*b*. The perturbation introduced by i_t causes a signal i_f to propagate in the *forward* direction, to which the loop responds with a *return* signal i_r. Let

$$T_i = -\frac{i_r}{i_f} \tag{8.15b}$$

It has been proved[4] that the loop gain T is such that

$$\frac{1}{1+T} = \frac{1}{1+T_v} + \frac{1}{1+T_i} \tag{8.16}$$

Solving for T we get

$$T = \frac{T_v T_i - 1}{T_v + T_i + 2} \tag{8.17a}$$

Alternatively, using Eqs. (8.15), we have

$$T = \frac{(v_r/v_f) \times (i_r/i_f) - 1}{2 - v_r/v_f - i_r/i_f} \tag{8.17b}$$

EXAMPLE 8.3. Suppose the circuit of Example 8.2 is implemented with a 741 op amp powered from ±10-V supplies. Find the phase margin using PSpice's 741 macromodel.

Solution. Lacking a dependent-source model for the op amp, we cannot use return-ratio analysis. Rather, we use the successive-injection test circuits of Fig. 8.14, and obtain the plots of Fig. 8.15. It is interesting to observe the way T_v and T_i contribute to the makeup of T. Using PSpice's cursor facility, we readily find $f_x = 283.6$ kHz and $\sphericalangle T(jf_x) = -122.7^\circ$, so $\phi_m = 180^\circ - 122.7^\circ = 57.3^\circ$.

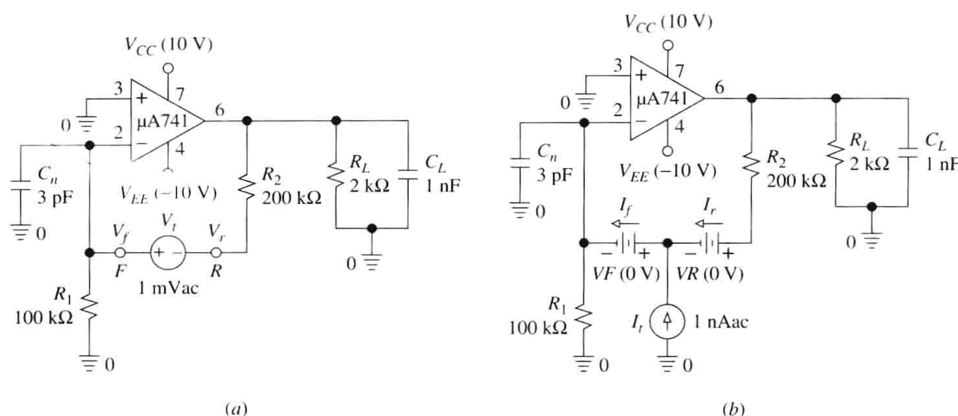


FIGURE 8.14
Finding the loop gain of the circuit of Example 8.3 via (a) voltage and (b) current injections.

Remark. One might be tempted to break the loop right at the op amp's output lead (pin #6) and use return-ratio analysis instead of double injections. This would be wrong because the test source would be driving C_L *directly*, thus preventing it from forming a pole with the op amp's internal resistance r_o. There is no question that the injection method is mandatory for this circuit!

Single-Injection Approximations

Since the loop gain $T(jf)$ is an intrinsic circuit characteristic, it must be independent of where we break the loop for our injections. If we break the loop of Fig. 8.14 at a different point, we will get the same curves for T, though those for T_v and T_i will generally vary with the point of injection. We wonder whether one injection point

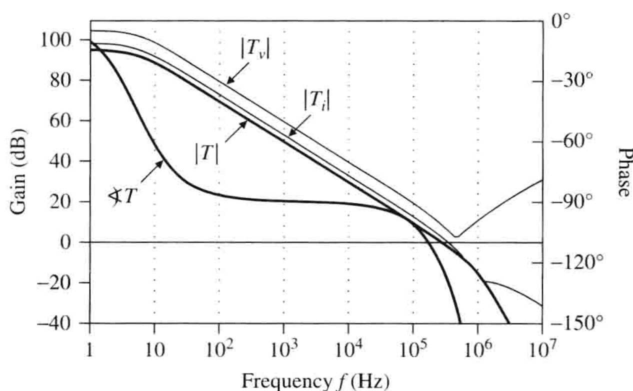


FIGURE 8.15
Bode plots for the circuit of Example 8.3. The $|T_v|$ and $|T_i|$ traces are plotted as `DB(-V(R)/V(F))` and as `DB(-I(VR)/I(VF))`, the $|T|$ and ∢T trace are plotted as `DB((V(R)*I(VR)-V(F)*I(VF))/(2*V(F)*I(VF)-V(R)*I(VF)-V(F)*I(VR)))` and as `P((V(R)*I(VR)-V(F)*I(VF))/(2*V(F)*I(VF)-V(R)*I(VF)-V(F)*I(VR)))`.

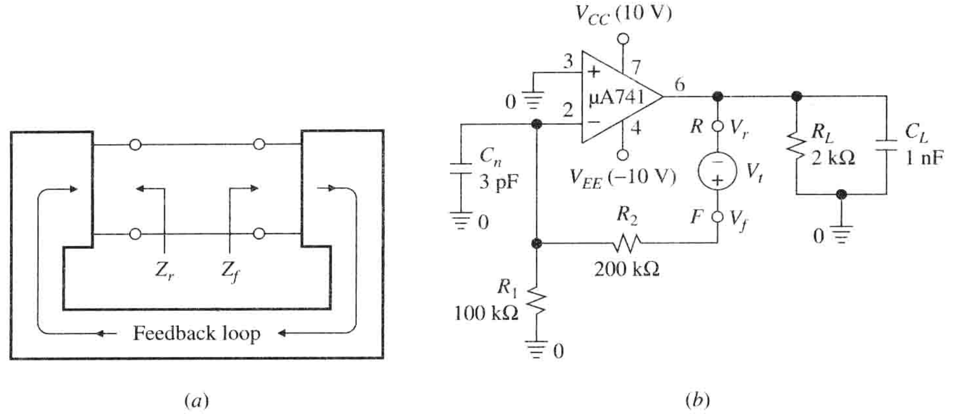


FIGURE 8.16
(*a*) The impedances Z_f and Z_r seen looking in the *forward* and *return* directions. (*b*) Single-injection approximation for the circuit of Example 8.3.

is preferable to another, and why. The answer is provided by the fact that the terms involving T_v and T_i satisfy the condition[4]

$$\frac{1+T_v}{1+T_i} = \frac{Z_r}{Z_f} \tag{8.18}$$

where Z_f and Z_r are the impedances seen looking in the *forward* and in the *return* directions relative to the point of signal injection (see Fig. 8.16*a*). According to Eq. (8.16) the terms $(1 + T_v)$ and $(1 + T_i)$ combine in the guise of resitances in parallel, so if one happens to be much larger than the other, the smaller will prevail and we can estimate T more quickly by limiting ourselves to just *one* signal injection, namely, the one resulting in the smaller of the two terms. Sure enough, if we break the loop of the circuit of Example 8.3 just upstream of R_2, where the condition $Z_f \gg Z_r$ holds, we have $(1 + T_v) \ll (1 + T_i)$, so we can use only a voltage injection and approximate $T \cong T_v = -V_r/V_f$. Indeed, running the PSpice circuit of Fig. 8.16*b* gives again $f_x = 283.6$ kHz and $\phi_m = 57.3°$, but via a single injection!

Feedthrough Considerations

To facilitate our introduction to stability, we have deliberately ignored the possibility of direct signal transmission from input to output, bypassing the loop. For a more accurate investigation we must refer to the more general block of Fig. 1.37, whose gain is, by Eq. (1.72),

$$A = \frac{A_{\text{ideal}}}{1 + 1/T} + \frac{a_{\text{ft}}}{1+T}$$

where a_{ft} is the feedthrough gain, which coincides with the value of A in the limit $a \to \infty$. Using Eq. (8.1), we manipulate A into the insightful form

$$A(jf) = A_{\text{ideal}} D_{\text{eff}}(jf) \tag{8.19a}$$

where the effective discrepancy function

$$D_{\text{eff}}(jf) = D(jf)\left[1 + \frac{a_{\text{ft}}(jf)}{T(jf)A_{\text{ideal}}}\right] \tag{8.19b}$$

represents a refinement of $D(jf)$ to account for feedthrough. As seen in the various examples of Chapter 1, if $T(jf)$ is sufficiently large to make $|T(jf)A_{\text{ideal}}| \gg |a_{ft}(jf)|$, the effect of feedthrough is negligible.[5] However, this may no longer be the case in the vicinity of the crossover frequency f_x, where $T(jf_x)$ drops to unity.

As an example, consider the I-V converter of Fig. 8.17, based on an op amp whose poles and dc gain have been adjusted for $\phi_m = 67.4°$, a slightly higher margin than that marking the onset of peaking, which is $\cos^{-1}\left(\sqrt{2}-1\right) = 65.5°$. Yet, the actual plot of $|A|$ reveals peaking even though loop-gain analysis predicts none. Evidently, signal feedthrough bolsters V_o in the vicinity of f_x in a way not accounted for by loop analysis alone. (The culprit in this particular example is the relatively high value of r_o compared to R, so chosen in order to cause substantial feedthrough. Feedthrough is evidenced also by the initial jump of -4 V of the transient response, occurring before the op amp starts to react.)

The interested reader is encouraged to consult the literature[6,7] for computer simulation techniques offering information about the closed-loop gain also in the presence of substantial feedthrough. Here, we take the viewpoint of using loop-gain analysis to come up with an initial stability estimate, and then fine-tuning the component values when testing the final circuit in the lab.

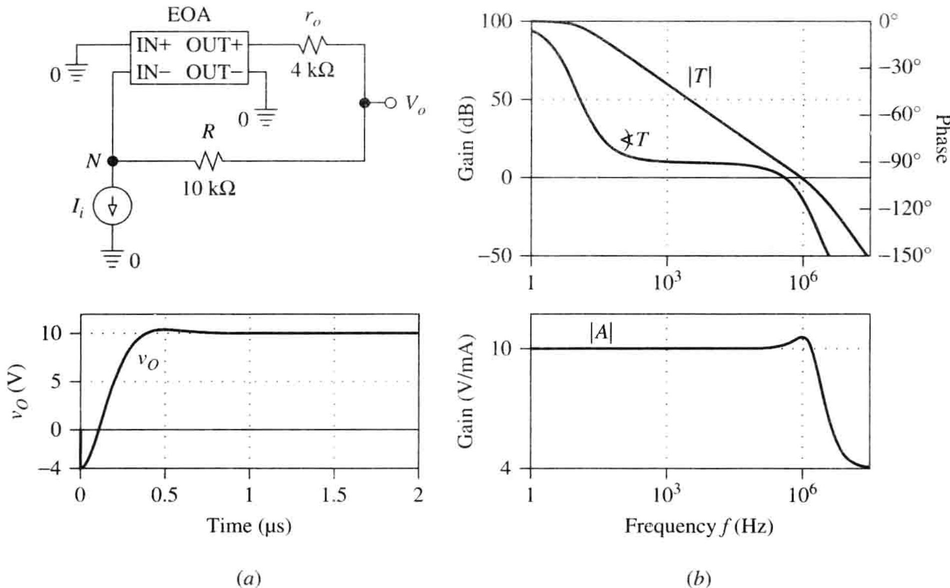


FIGURE 8.17
(*a*) PSpice I-V converter (top), and transient response to a 1-mA step (bottom). (*b*) Loop gain T (top), and closed-loop gain $A = V_o/I_i$ (bottom).

8.3 FREQUENCY COMPENSATION OF OP AMPS

As a signal travels around a feedback loop, it experiences a variety of delays, first in propagating through the transistors of the op amp, then through the reactive elements (parasitic as well as intentional, if any) of the surrounding circuitry. If the cumulative delay is such that $\measuredangle T(jf_x) \le -180^\circ$, the circuit will oscillate, so we need to modify $T(jf)$ to guarantee an adequate phase margin ϕ_m, a modification known as *frequency compensation*. Since $T = a\beta$, frequency compensation requires that we modify $a(jf)$, or $\beta(jf)$, or both.

We begin by investigating how to modify $a(jf)$ for the case of *frequency-independent* feedback. Although this is usually the task of the IC designer, the user needs to be familiar with the various compensation schemes in use as they may influence the choice of the device for a given application. Typically, the IC designer strives to guarantee a given ϕ_m for the hardest-to-compensate configuration, namely, the voltage-follower, for which $\beta = 1$. In this particular case we have $T(jf) = a(jf) \times 1 = a(jf)$, that is, $T(jf)$ *coincides* with $a(jf)$. As a vehicle, let us use the generalized op amp model of Fig. 8.18, consisting of two transconductance stages with dc gains $a_{10} = -g_1R_1$ and $a_{20} = -g_2R_2$, followed by a unity-gain voltage stage, so the overall dc gain is $a_0 = a_{10}a_{20}1$. The three R-C networks establish three pole frequencies at $f_k = 1/(2\pi R_kC_k)$, $k = 1, 2$, and 3. With the component values shown, we have

$$a_0 = (-200)(-500)1 = 10^5 \text{ V/V} \quad f_1 = 1 \text{ kHz} \quad f_2 = 100 \text{ kHz} \quad f_3 = 10 \text{ MHz}$$

(These parameters have deliberately been chosen to match those of Fig. 8.4, so we can recycle already familiar knowledge.)

Dominant-Pole Compensation

A popular compensation technique involves *lowering* the first pole frequency f_1 to a new value $f_{1(\text{new})}$ such that the compensated response is dominated by this lone pole *all the way* up to the crossover frequency, which for $\beta = 1$ is the *transition frequency* f_t. We can then write $\measuredangle T(jf_t) = -90^\circ + \phi_{t(\text{HOR})}$, where -90° is the phase shift due to $f_{1(\text{new})}$, and $\phi_{t(\text{HOR})}$ is the *combined phase shift* due to the *higher-order roots* (poles and possibly zeros) at f_t. (For instance, the circuit of Example 8.3 has $\phi_{t(\text{HOR})} = 90 - 122.7 = -32.7^\circ$.) The phase margin after compensation is

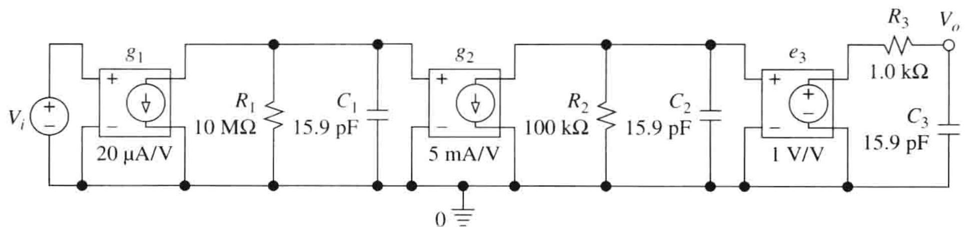


FIGURE 8.18
Generalized three-pole op amp model for the investigation of frequency compensation.

$\phi_m = 180^\circ + \measuredangle T(jf_t) = 180^\circ - 90^\circ + \phi_{t(\text{HOR})}$, that is,

$$\phi_m = 90^\circ + \phi_{t(\text{HOR})} \tag{8.20}$$

By making $f_{1(\text{new})}$ sufficiently low, we can keep $\phi_{t(\text{HOR})}$ as small as needed. For instance, for $\phi_m = 60^\circ$ we need $\phi_{x(\text{HOR})} = -30^\circ$.

A simple way to lower f_1 is to deliberately *increase* the capacitance at the node responsible for f_1 itself. In the three-pole amplifier of Fig. 8.18 we simply add a *shunt capacitance* C_c in parallel with C_1 to lower the first pole frequency from $f_1 = 1/(2\pi R_1 C_1)$ to

$$f_{1(\text{new})} = \frac{1}{2\pi R_1(C_1 + C_c)} \tag{8.21}$$

To estimate the required value of $f_{1(\text{new})}$, note that a dominant-pole response exhibits a *constant gain-band width product*, GBP $= a_0 \times f_{1(\text{new})} = 1 \times f_t$. So, we first select f_t on the basis of the desired phase margin, then we estimate the dominant pole frequency as

$$f_{1(\text{new})} = \frac{f_t}{a_0} \tag{8.22a}$$

Finally, we substitute into Eq. (8.21) and solve for the desired C_c.

EXAMPLE 8.4. (*a*) Find the capacitance C_c that, placed in parallel with C_1 in Fig. 8.18, will ensure $\phi_m \cong 45^\circ$ for operation with $\beta = 1$. (*b*) Verify with PSpice and comment on your results. (*c*) Repeat, but for $\phi_m \cong 60^\circ$.

Solution.

(*a*) For $\phi_m \cong 45^\circ$ we need to impose $f_t = f_{-135^\circ}$. Inspection of Fig. 8.4 reveals that $f_{-135^\circ} = f_2 = 100$ kHz, so Eq. (8.22*a*) gives $f_{1(\text{new})} = 10^5/10^5 = 1$ Hz. Plugging into Eq. (8.21) we get

$$1 = \frac{1}{2\pi 10^7(15.9 \times 10^{-12} + C_c)}$$

whose solution is $C_c = 999C_1 \cong 15.9$ nF ($\gg C_1$).

(*b*) Using the PSpice circuit of Fig. 8.19, we generate the plots of Fig. 8.20. Cursor measurements give $f_t = 78.6$ kHz and $\measuredangle a(jf_t) = -128.6^\circ$, so $\phi_m = 180 - 128.6 = 51.4^\circ$, in fair agreement with the target values of $f_t = 100$ kHz and $\phi_m = 45^\circ$. We note that f_t is much lower than the crossover frequency of the uncompensated response, which is near 3 MHz. However, Fig. 8.4 reveals that the phase shift there is about -195°, indicating that with $\beta = 1$ the circuit would be unstable. Evidently, the drastic bandwidth reduction is the price we are paying for the sake of stability!

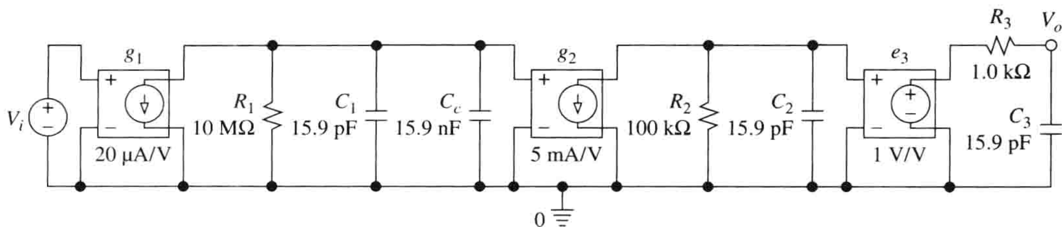


FIGURE 8.19
Dominant-pole compensation for $\phi_m \cong 45^\circ$ with $\beta = 1$ using shunt capacitance C_c.

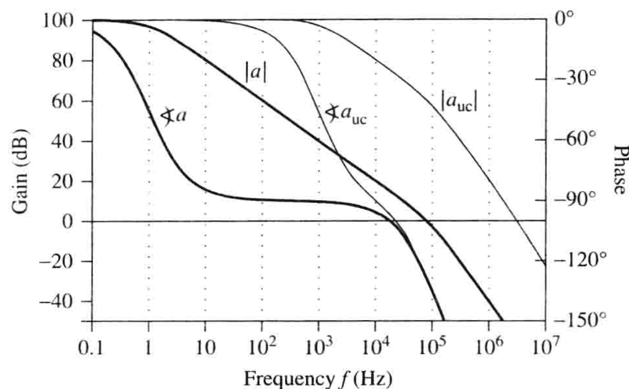


FIGURE 8.20
Compensated (a) and uncompensated (a_{uc}) responses of the circuit of Fig. 8.19.

(c) For $\phi_m = 60^\circ$ we need to impose $f_t = f_{-120^\circ}$ (= 59.2 kHz, as per Example 8.1). Recalculating as above we now get the more conservative values $f_{1(\text{new})} = 0.592$ Hz and $C_c \cong 26.9$ nF. Rerunning PSpice with this new value of C_c gives $f_t = 52.4$ kHz and $\measuredangle a(jf_t) = -118^\circ$, so now $\phi_m = 62^\circ$.

EXAMPLE 8.5. Find the capacitance C_c that, placed in parallel with C_1 in Fig. 8.18, will ensure $\phi_m \cong 60^\circ$ for operation with $\beta = 0.1$. Hence, use PSpice to plot both the open-loop and the closed-loop gains.

Solution. So long as β is frequency independent, $T(jf)$ will have the *same* poles as $a(jf)$. However, we now have GBP $= a_0 \times f_{1(\text{new})} = (1/\beta) \times f_x$, so Eq. (8.22$a$) changes as

$$f_{1(\text{new})} = \frac{f_x}{\beta a_0} \tag{8.22b}$$

For $\phi_m = 60^\circ$ we get $f_{1(\text{new})} = (59.2 \times 10^3)/(0.1 \times 10^5) = 5.92$ Hz. Plugging into Eq. (8.21) gives $C_c \cong 2.67$ nF. Using the PSpice circuit of Fig. 8.21, we generate the plots of Fig. 8.22, where we visualize the loop gain $|T|$ as the difference between the $|a|$ and $1/\beta$ curves. Using the cursor, we find that the two curves intersect at $f_x = 52.5$ kHz,

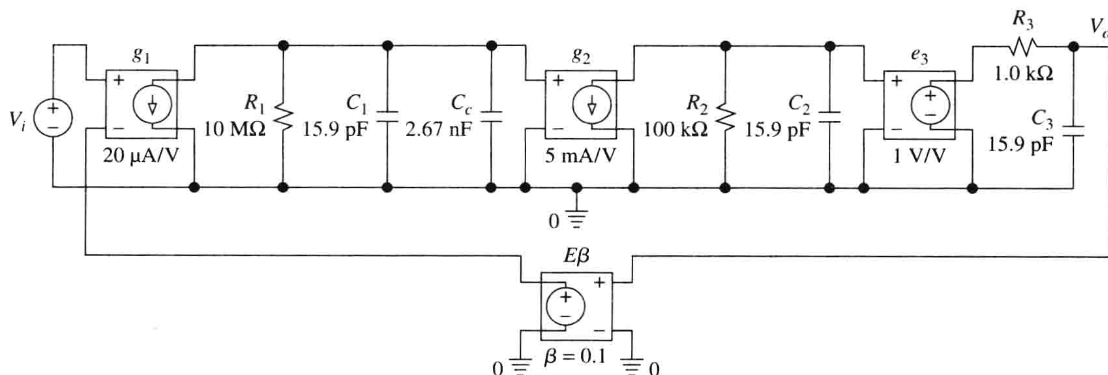


FIGURE 8.21
Compensating the amplifier of Fig. 8.18 for $\phi_m = 60^\circ$ with $\beta = 0.1$ V/V.

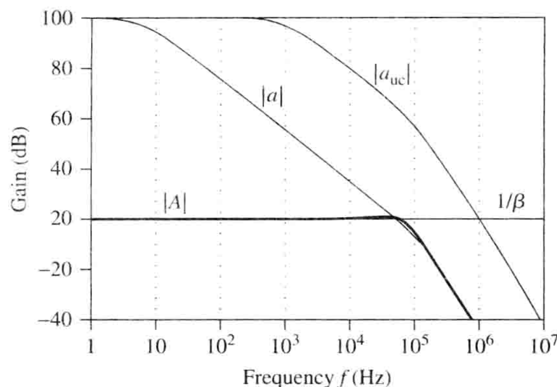


FIGURE 8.22
Open-loop a, closed-loop A, and uncompensated a_{uc} gains of the circuit of Fig. 8.21.

where $\measuredangle a(jf_x) = -118°$, so $\phi_m = 62°$ (note the slight peaking, in accordance with Fig. 8.7). We observe that as we increase $1/\beta$, compensation becomes less conservative. The reader may wish to verify that without C_c, the circuit would be close to oscillatory, as indicated by Figs. 8.5 and 8.6.

It must be pointed out that the model of Fig. 8.18 is based on stages that are isolated from each other, so changing the capacitance of one stage does not affect the other stages. In an actual amplifier, adding capacitance to change one pole is likely to have some effect also on the other poles, so the compensation schemes based on the above model provide just a starting point. To meet the desired ϕ_m, we may have to do some tweaking, a task greatly facilitated by the visual feedback attainable via PSpice.

Pole-Zero Compensation

It is possible to increase the GBP by the mere artifice of inserting a suitable resistance $R_c (\ll R_1)$ in series with C_c. It can be proved (see Problem 8.28) that the presence of R_c introduces a *left-half-plane zero* (LHPZ), whose phase lead can be exploited on purpose to compensate for a higher f_t. A common approach is to make this zero coincide with the existing second pole so as to achieve a *pole-zero cancellation* that raises phase by about 90° at the pole-zero itself.

Applying physical insight to Fig. P8.28, we can state the following: (*a*) At low frequencies C_c's impedance is much larger than R_c, so R_c acts as a short by comparison, and Eq. (8.21) still holds. (*b*) As we raise f, C_c's impedance decreases until its magnitude equals R_c. This occurs at the frequency f_z such that $1/|2\pi j f_z C_c)| = R_c$, or $f_z = 1/(2\pi R_c C_c)$. (*c*) For $f \gg f_z$, C_c acts as a short compared to R_c, leaving C_1 to form an additional pole with the combination $R_1 \| R_c (\cong R_c$, since $R_c \ll R_1)$. In summary, the presence of R_c leaves $f_{1(\text{new})}$ unchanged, but introduces a *zero frequency* f_z as well as a *new pole frequency* f_4,

$$f_{1(\text{new})} = \frac{1}{2\pi R_1(C_1 + C_c)} \qquad f_z = \frac{1}{2\pi R_c C_c} \qquad f_4 \cong \frac{1}{2\pi R_c C_1} \tag{8.23}$$

EXAMPLE 8.6. Find R_c and C_c to compensate the circuit of Fig. 8.18 for $\phi_m \cong 45°$ with $\beta = 1$. Verify with PSpice.

Solution. After the pole-zero cancellation, $a(jf)$ is left with three pole frequencies: $f_{1(\text{new})}$, f_3, and $f_4(f_4 \gg f_{1(\text{new})})$. Let us momentarily ignore f_3 and impose $f_t = f_4$ for $\phi_m \cong 45°$. Letting $f_{1(\text{new})} = f_t/a_0$ and $f_z = f_2$, we have

$$\frac{1}{2\pi R_1(C_1 + C_c)} = \frac{1}{a_0}\frac{1}{2\pi R_c C_1} \qquad \frac{1}{2\pi R_c C_c} = f_2$$

Plugging in the known data ($R_1 = 10$ MΩ, $C_1 = 15.9$ pF, $a_0 = 10^5$ V/V, and $f_2 = 100$ kHz), we get two equations in the unknowns R_c and C_c, which we easily solve to get $R_c = 3.2$ kΩ and $C_c = 496$ pF. Running PSpice with these values results in $f_t = 2.42$ MHz and $\measuredangle a(jf_t) = -140.4°$, so $\phi_m = 39.6°$. This is less than the target margin of 45° because we have ignored f_3. To increase ϕ_m we need to raise C_c a bit while lowering R_c in proportion in order to maintain pole-zero cancellation. This causes $f_{1(\text{new})}$ and f_4 to split a bit further apart. After a few iterations aided by visual feedback from PSpice, we find that raising C_c to 560 pF while reducing R_c to 2.84 kΩ gives $f_t = 2.28$ MHz and $\measuredangle a(jf_t) = -135°$, so $\phi_m = 45°$. The final circuit is shown in Fig. 8.23, and its response in Fig. 8.24, which you can compare with Fig. 8.20 to appreciate the wider bandwidth afforded by pole-zero cancellation.

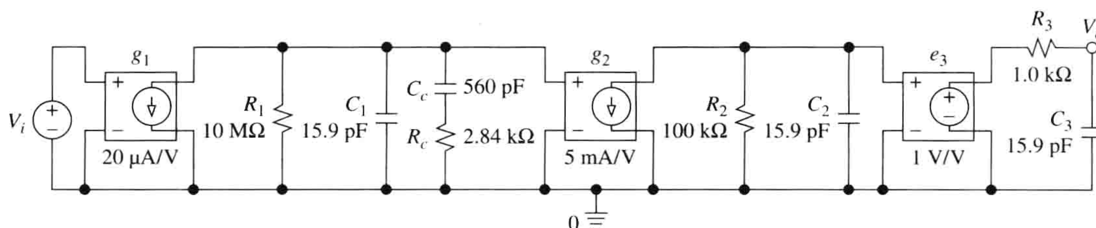


FIGURE 8.23
Pole-zero compensation for $\phi_m \cong 45°$ with $\beta = 1$.

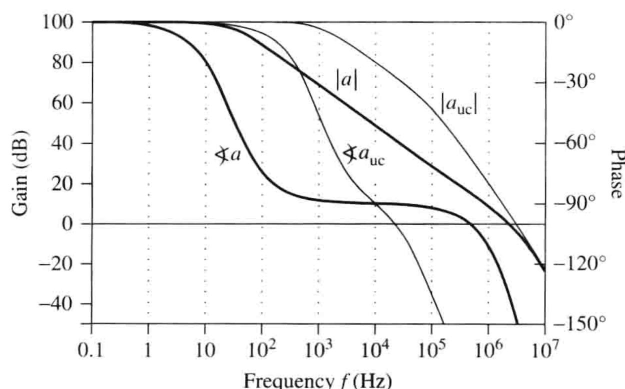


FIGURE 8.24
Compensated (a) and uncompensated (a_{uc}) responses of the circuit of Fig. 8.23.

Miller Compensation

Nowadays the majority of op amps come with frequency compensation on chip, a trend pioneered by the 741 op amp in the late 1960s. The shunt and pole-zero schemes discussed above, though pedagogically insightful, do not lend themselves to on-chip fabrication because of the large C_c required (capacitors of more than a few tens of picofarads would take up too much chip area). A clever way around this limitation is to start out with a small enough capacitance C_c that can be fabricated on chip, and then use the *Miller effect* to make it appear as large as needed for dominant-pole compensation. We shall see that two additional benefits accrue from this scheme, namely, *pole splitting* and *higher slew rates*.

In order to focus on the essentials of Miller compensation, let us restrict ourselves to the first two stages, shown in Fig. 8.25. In the absence of C_c, we know that the circuit exhibits the pole frequencies

$$f_1 = \frac{1}{2\pi R_1 C_1} \qquad f_2 = \frac{1}{2\pi R_2 C_2} \tag{8.24}$$

With C_c in place, a detailed ac analysis (see Problem 8.31) yields

$$\frac{V_2}{V_i} = a_0 \frac{1 - jf/f_z}{\left(1 + jf/f_{1(\text{new})}\right)\left(1 + jf/f_{2(\text{new})}\right)} \tag{8.25}$$

where

$$a_0 = g_1 R_1 g_2 R_2 \qquad f_z = \frac{g_2}{2\pi C_c} \tag{8.26a}$$

$$f_{1(\text{new})} \cong \frac{f_1}{(g_2 R_2 C_c)/C_1} \qquad f_{2(\text{new})} \cong \frac{(g_2 R_2 C_c) f_2}{C_1 + C_c(1 + C_1/C_2)} \tag{8.26b}$$

Since f_1 gets *divided* and f_2 gets *multiplied* by the common term $(g_{m2} R_2 C_c)$, the presence of C_c *lowers* the first pole and *raises* the second pole, a phenomenon known as *pole splitting*. Depicted in Fig. 8.26 for $\phi_m \cong 45°$, pole splitting is highly welcome because the upshift in $f_{2(\text{new})}$ eases the amount of downshift required of $f_{1(\text{new})}$, now the dominant pole, thus allowing for a wider GBP than with shunt-capacitance compensation. Writing GBP $= a_0 \times f_{1(\text{new})}$ and using Eqs. (8.24) and (8.26), we have

$$\text{GBP} \cong \frac{g_1}{2\pi C_c} \tag{8.27}$$

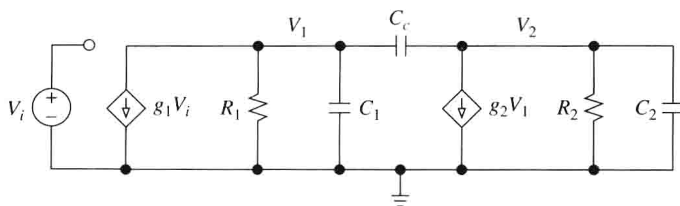


FIGURE 8.25
Two-pole amplifier with Miller compensation.

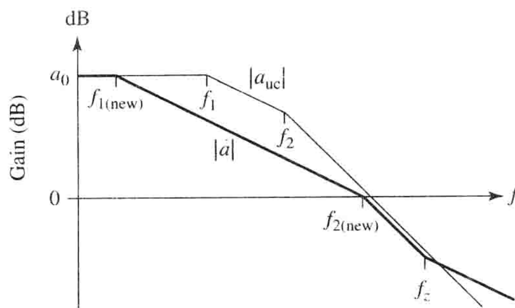


FIGURE 8.26
Miller compensation and pole splitting for the two-pole amplifier of Fig. 8.25.

It is instructive to manipulate the dominant pole into the alternative form

$$f_{1(\text{new})} \cong \frac{1}{2\pi R_1[(g_2 R_2)C_c]} \tag{8.28}$$

because it shows that thanks to the Miller effect, the (small) capacitance C_c gets *multiplied* by the (large) second-stage gain ($g_2 R_2 = 500$ in our example) to form, with R_1, the (low-frequency) dominant pole. Yet, during large transients, it is C_c (not the multiplied C_c!) that is responsible for the slew rate, as already seen in Eq. (6.33) for the 741 op amp.

We observe that Miller compensation, besides relocating the pole pair, establishes also a right-half-plane (RHP) zero, which produces *phase lag*, possibly eroding the phase margin. Physically, when the current injected into node V_2 via C_c equals that removed by the $g_2 V_1$ source, R_2 is left with no current, so V_2 goes to *zero*, as shown in Fig. 8.27*a*. This occurs at the frequency f_z such that $(V_1 - 0)/|1/(2\pi j f_z C_c)| = g_2 V_1$, that is, for f_z as given in Eq. (8.26*a*). For $f > f_z$, the incoming current via C_c prevails over that removed by the $g_2 V_1$ source, resulting in *polarity reversal*, so feedback turns from negative to positive.

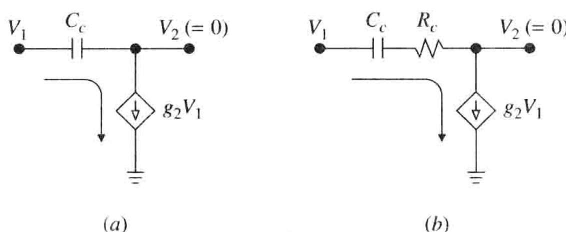


FIGURE 8.27
Partial circuits to visualize (*a*) the formation of the RHP zero, and (*b*) zero control via R_c.

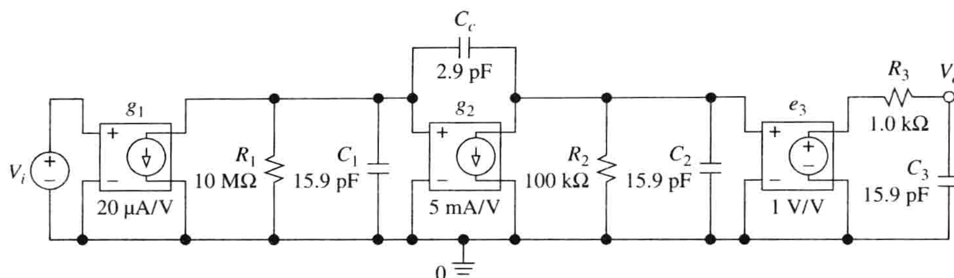


FIGURE 8.28
Miller compensation for $\phi_m \cong 75°$ with $\beta = 1$.

EXAMPLE 8.7. Find C_c to Miller-compensate the circuit of Fig. 8.18 for $\phi_m \cong 75°$ with $\beta = 1$. Verify with PSpice.

Solution. By Eq. (8.20) we need $\phi_{t(\text{HOR})} = -15°$. Assume for a moment that $\phi_{t(\text{HOR})}$ is due exclusively to $f_{2(\text{new})}$. Then, f_t must be such that $-15° = -\tan^{-1}(f_t/f_{2(\text{new})})$, or $f_t = 0.268 f_{2(\text{new})}$. Imposing $f_{1(\text{new})} = f_t/a_0$ gives

$$\frac{f_1}{(g_2 R_2 C_c)/C_1} = \frac{0.268}{10^5} \frac{(g_2 R_2 C_c) f_2}{C_1 + C_c(1 + C_1/C_2)}$$

Plugging the known data and solving gives $C_c = 2.2$ pF. Running the PSpice circuit of Fig. 8.18 with $C_c = 2.2$ pF gives $\phi_m = 68°$, so we need to increase C_c a bit because of the phase lag by the higher-order roots. Empirically we find that $C_c = 2.9$ pF gives $f_t = 1.06$ MHz and $\phi_m = 74.8°$. The final circuit, shown in Fig. 8.28, gives the response in Fig. 8.29 (note that because of the $-90°$ shift by the RHP zero, the high-frequency asymptote of phase changes from $-270°$ to $-360°$).

RHP Zero Control

The RHP zero frequency in Example 8.7 is $f_z = 5 \times 10^{-3}/(2\pi \times 2.9 \times 10^{-12}) = 274$ MHz, so much higher than f_t that its effect on ϕ_m is immaterial. This, however,

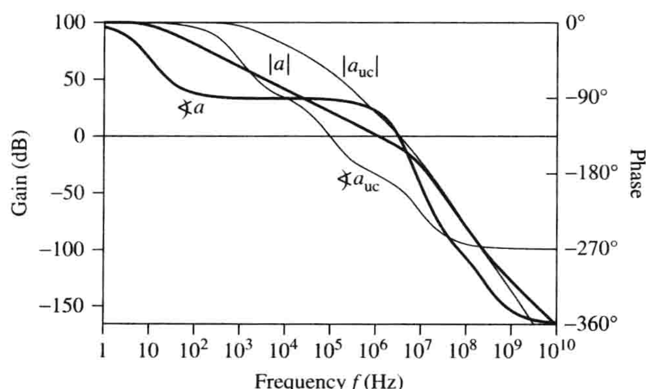


FIGURE 8.29
Compensated (a) and uncompensated (a_{uc}) responses of the circuit of Fig. 8.28.

is not always the case (the two-stage CMOS op amp is a notorious example in which the RHP zero does matter[8,9]). To investigate further, use Eq. (8.26) to express f_z in the alternative form

$$f_z \cong \frac{g_2}{g_1}\text{GBP} \tag{8.29}$$

where $\text{GBP} = a_0 \times f_{1(\text{new})}$. So long as $g_2 \gg g_1$, we have $f_z \gg \text{GBP}$, so $f_t \cong \text{GBP}$ and f_z has negligible effect. Not so if g_2 is close to, or even less than g_1, for then we have a two-fold reduction in ϕ_m, first because the +20-dB slope change due to f_z raises f_t, second because of the phase lag due to f_z itself.

A clever way around this drawback is to insert a suitable resistance R_c in series with C_c, as shown in Fig. 8.27*b*. At low frequencies, where C_c's impedance is much larger than R_c, we can regard R_c as a short by comparison, so Eq. (8.26*b*) still holds for the pole pair. However, the presence of R_c changes the location of the transmission zero, now occurring at the complex frequency s_z such that $(V_1 - 0)/[1/(s_z C_c) + R_c] = g_2 V_1$. Solving, we get the *s-plane zero*

$$s_z = \frac{1}{(1/g_2 - R_c)\, C_c} \tag{8.30a}$$

so the zero frequency in the numerator of Eq. (8.25) takes on the modified form

$$f_{z(\text{new})} = \frac{1}{2\pi(1/g_2 - R_c)C_c} \tag{8.30b}$$

The presence of R_c $(0 < R_c < 1/g_2)$ reduces the denominator, raising $f_{z(\text{new})}$ and pushing the phase lag away from f_x. For $R_c = 1/g_2$ we have $f_{z(\text{new})} \to \infty$, so the numerator of Eq. (8.25) becomes unity. Raising R_c further ($R_c > 1/g_2$) *changes the polarity* of $f_{z(\text{new})}$, resulting in a *left-half-plane* (LHP) zero. This is highly desirable because it produces phase *lead* (as opposed to phase *lag* by an RHP zero), thus boosting ϕ_m.

EXAMPLE 8.8. (*a*) Use PSpice to find the phase margin of the two-stage op amp of Fig. 8.30 in the absence of compensation, and comment. (*b*) Show how to calculate the values of C_c and R_c given in the figure, and use PSpice to measure the actual values of f_t and ϕ_m. (*c*) Discuss the effect of letting $R_c = 0$, and verify with PSpice. (*d*) How would you change R_c to raise ϕ_m to 60°?

Solution.

(*a*) Running PSpice without the C_c-R_c network we find $f_t = 223$ MHz and $\sphericalangle a(jf_t) = -177°$, so $\phi_m = 3°$, indicating a circuit in dire need of compensation.

(*b*) By inspection, $a_0 = (-100) \times (-50) = 5{,}000$ V/V, $f_1 = 1$ MHz, and $f_2 = 10$ MHz. To achieve $f_{z(\text{new})} \to \infty$, impose $R_c = 1/g_2 = 1/10^{-3} = 1\ \text{k}\Omega$. Then, for $\phi_m \cong 45°$ impose $f_t = f_{2(\text{new})}$. Plugging the data into Eq. (8.26*b*), and then into Eq. (8.22*a*), and solving, we get $C_c = 2.144$ pF. Running PSpice with the C_c-R_c network shown in Fig. 8.30, we get the response of Fig. 8.31, for which $f_t = 58.7$ MHz and $\sphericalangle a(jf_t) = -129.4°$, so $\phi_m = 50.6°$.

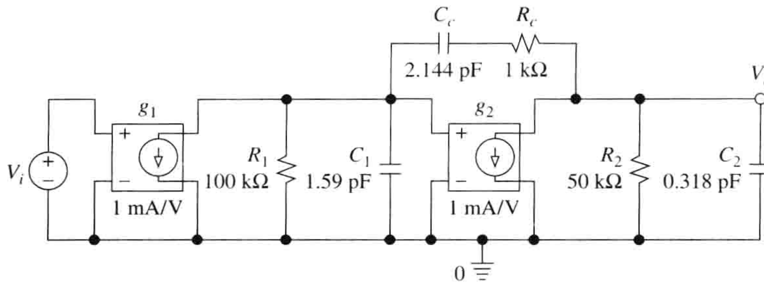


FIGURE 8.30
Miller compensation for $\phi_m \cong 45°$ with $\beta = 1$, under the constraint $f_{z(\text{new})} \to \infty$.

(*c*) For $R_c = 0$, use Eqs. (8.26) to calculate $f_z \cong 74.2$ MHz and $f_{2(\text{new})} \cong 74.2$ MHz. These frequencies are identical in value, but since the zero is in the RHP and the pole in the LHP, the zero and pole are *complex-conjugate* of each other. As we use Eq. (8.25) to calculate magnitude, the terms involving f_z and $f_{2(\text{new})}$ cancel each other out, leaving *only* the term involving $f_{1(\text{new})}$. However, as we calculate phase, *both* f_z and $f_{2(\text{new})}$ contribute *phase lag*, drastically reducing ϕ_m. This is confirmed by the thin traces of Fig. 8.31, obtained by rerunning PSpice with $R_c = 0$. Cursor measurements now give $f_t = 72.4$ MHz and $\measuredangle a(jf_t) = -177.3°$, so $\phi_m = 2.7°$. No doubt that R_c plays a vital role! (Beware that even though the situation is deceptively similar to that of $|1/\beta_4|$ in Fig. 8.10, ROC reasoning is not applicable here because this is not a minimum-phase circuit!)

(*d*) Raising R_c above 1 kΩ moves the zero to the LHP, thus introducing phase lead (and turning the circuit into a minimum-phase type!). Tweaking with R_c, we find that raising it to $R_c = 1.25$ kΩ yields $f_t = 60$ MHz and $\measuredangle a(jf_t) = -119.3°$, so $\phi_m = 60.7°$.

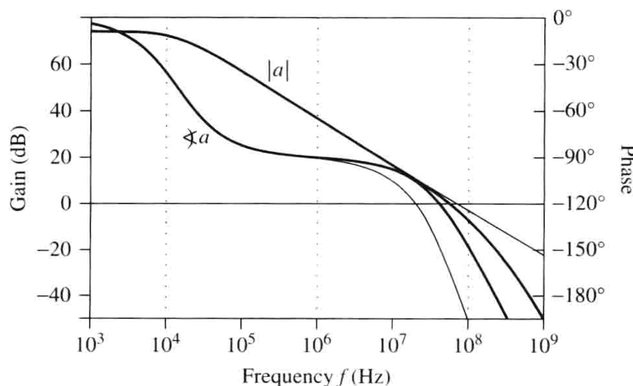


FIGURE 8.31
Frequency response of the circuit of Fig. 8.30 (thin curves show the response with $R_c = 0$).

Feedforward Compensation

In a multistage amplifier there is usually one stage that acts as a bandwidth bottleneck by contributing substantial phase lag. Feedforward compensation creates a high-frequency bypass around this bottleneck stage in order to reduce its phase lag contribution in the vicinity of the crossover frequency. The principle is illustrated in Fig. 8.32 (top) for the case of an inverting amplifier implemented with a two-stage op amp having pole frequencies of 1 kHz and 100 kHz, respectively. Clearly, the bottleneck stage is the first stage, and the bypass function is provided by the compensation capacitance C_{ff}.

For a closer look, let us plot T using the circuit at the bottom of Fig. 8.32, where the point of injection has been chosen so that a single injection suffices. It is apparent from Fig. 8.33 that without C_{ff} the rate of closure (ROC) of the $|T_{uc}|$ curve with the 0-db curve approaches 40 dB/dec, indicating a circuit on the verge of oscillation. The presence of C_{ff}, while lowering the bottleneck pole frequency from $1/(2\pi R_1 C_1)$ to about $1/[2\pi R_1(C_1 + C_{ff})]$, establishes also a zero frequency somewhere between 100 kHz and 1 MHz, whose effect is to cut the ROC almost in half. In fact, cursor measurements give $f_x = 2.56$ MHz and $\measuredangle T(jf_x) = -103.4°$, so $\phi_m = 76.6°$. The slight amount of peaking and overshoot appearing in Fig. 8.34 is due to the fact that the compensated open-loop response includes a *pole-zero doublet*. The step response consists[17] of two exponential transients acting in concert: a fast exponential transient leading to the overshoot and a slower exponential transient, also known as a *long*

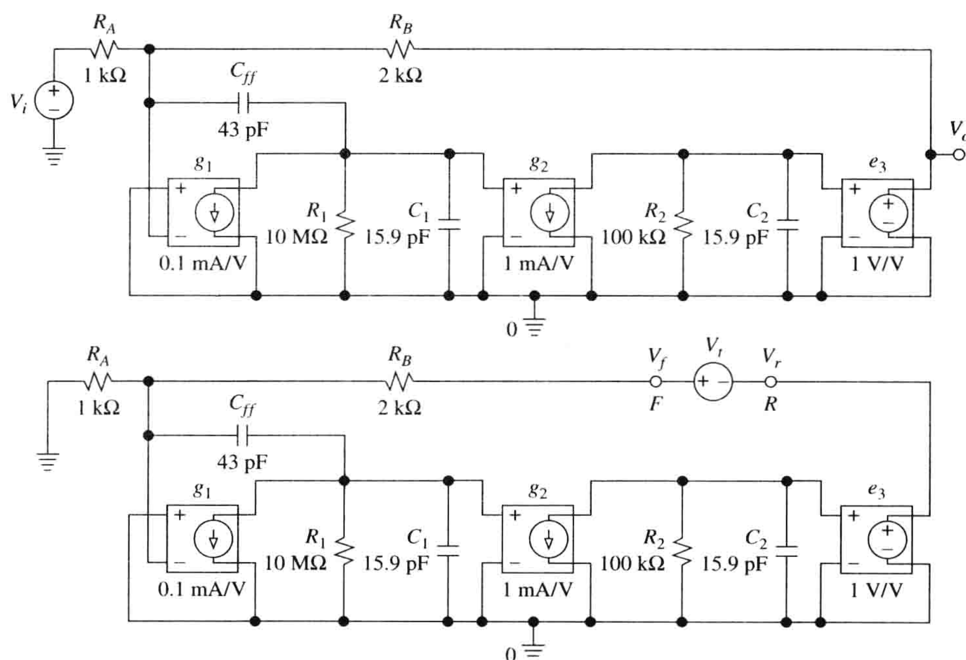


FIGURE 8.32
Feedforward compensation for an inverting amplifier (top), and voltage-injection circuit to plot T (bottom).

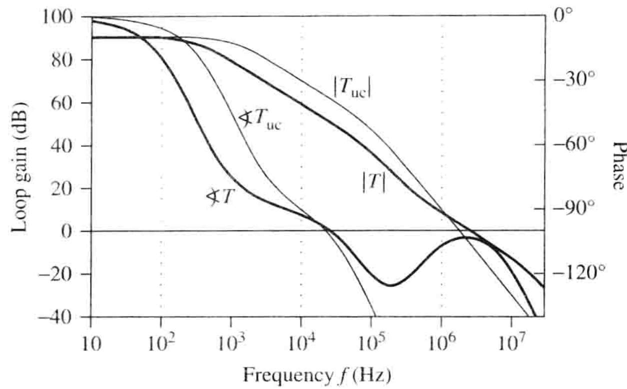


FIGURE 8.33
Loop gain of the circuit of Fig. 8.32 (T_{uc} refers to the uncompensated case).

tail, pulling back to the steady-state value of −2 V. The long tail tends to lengthen the settling time significantly,[17] and this may be an issue in certain applications.

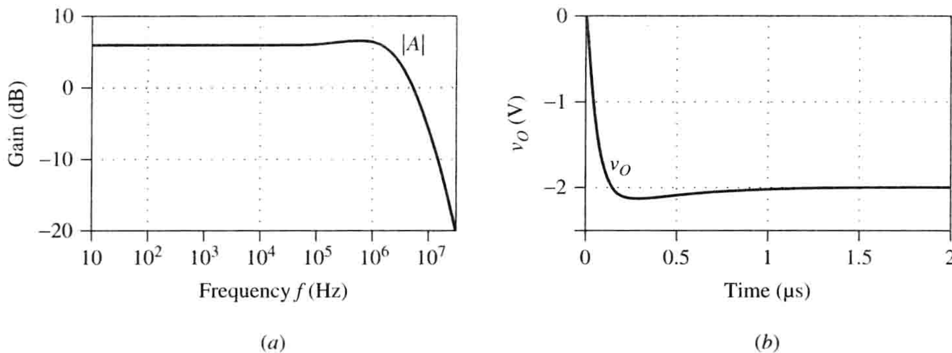


FIGURE 8.34
Closed-loop responses of the circuit of Fig. 8.32 (top): (*a*) frequency response and (*b*) transient response to a +1-V input step.

Three Representative Examples

Figure 8.35 shows three examples of the frequency compensation techniques discussed above.

The 741 circuit schematic of Fig. 5A.2 reveals Miller compensation via the 30-pF capacitor in the feedback path of the Darlington pair forming the second stage (the compensation subcircuit is repeated in Fig. 8.35*a* for convenience). The second stage has a gain of about −500 V/V, indicating that once reflected to the input of the second stage, C_c appears as a capacitance of about 500× (30 pF) = 15 nF.

Figure 8.35*b* shows the compensation subcircuit of the two-stage CMOS op amp of Fig. 5.4*a*. Here too we have Miller compensation. However, because of the notoriously low transconductance of MOSFETs (g_{m5} in this case), the zero

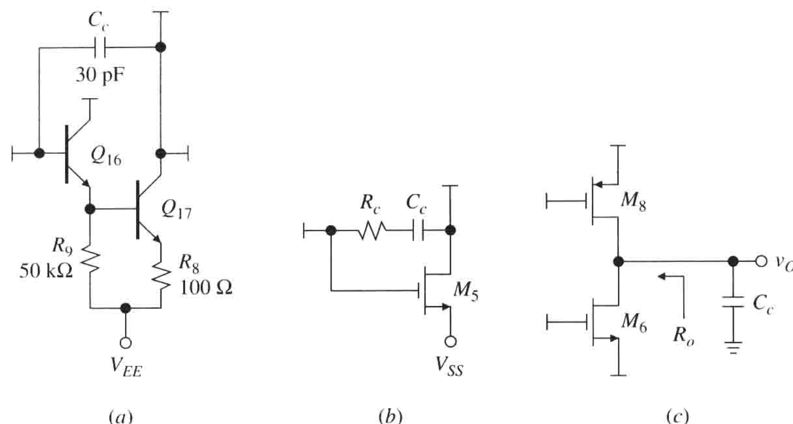


FIGURE 8.35
Subcircuits showing the frequency compensation of (*a*) the 741 bipolar op amp, (*b*) the two-stage CMOS op amp, and (*c*) the folded-cascode CMOS op amp.

frequency introduced by Miller compensation is low enough to erode the phase margin significantly. So, including a series resistance $R_c = 1/g_{m5}$ moves the zero out of the way to infinity.

Figure 8.35*c* shows the compensation subcircuit of the folded-cascode CMOS op amp of Fig. 5.4*b*. A distinguishing attribute of this topology is that all nodes except for the output node are low-resistance nodes, so the pole formed by each such node with its own stray capacitance is a high-frequency pole. Consequently, the overall response is dominated by the output pole $f_b = 1/(2\pi R_o C_c)$. Given that the dc gain is $a_0 = g_{m1} R_o$, it follows that GBP $= a_0 f_b = g_{m1}/(2\pi C_c)$. Adding capacitance at the output, far from destabilizing the op amp, will actually make ϕ_m approach 90°, if at the price of a lower GBP. For this reason folded-cascode op amps are particularly suited to applications that require driving capacitive loads, as in switched-capacitor filters.

8.4 OP AMPS CIRCUITS WITH A FEEDBACK POLE

Nowadays most op amps come with on-chip frequency compensation. Also referred to as *internally compensated* op amps, they use schemes of the type of Section 8.3 to provide, in frequency-independent feedback operation, a phase margin of the type

$$\phi_m = 90^\circ + \phi_{x(\text{HOR})}$$

where $\phi_{x(\text{HOR})}$ is the combined *phase lag* due to higher roots (HOR) at the crossover frequency f_x. Most op amps are compensated for operation up to $\beta = 1$, in which case f_x is the unity-gain frequency f_t (these op amps are also said to be *unity-gain stable*). To allow for faster closed-loop dynamics, some op amps are compensated for operation up to some $\beta_{(\text{max})} < 1$, such as $\beta_{(\text{max})} = 0.2$, in which case f_x is the crossover frequency between the $1/\beta_{(\text{max})}$ and the $|a(jf)|$ curves. Aptly called *de-compensated* op amps, they are intended for closed-loop gains of $1/\beta_{(\text{max})}$ or higher.

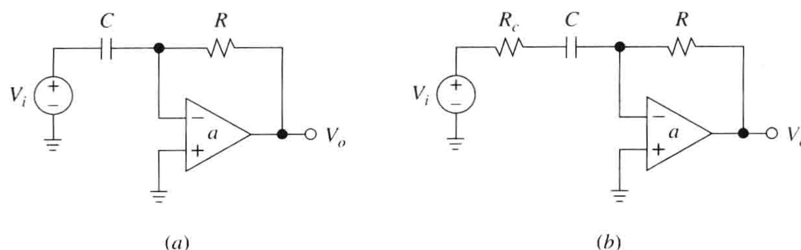


FIGURE 8.36
The differentiator circuit: (*a*) uncompensated, and (*b*) compensated by means of R_c.

If the feedback network contains reactive elements, whether intentional or parasitic, ϕ_m may get so low as to require that the user alter the feedback network to restore ϕ_m to the desired value. We are particularly concerned about the presence of poles in the feedback path because their phase lag tends to destabilize the loop. Example 8.2 has already exposed us to two feedback-pole types: that due to capacitance at the inverting input, and that due to capacitance at the output. We now wish to investigate the two cases separately. To focus on the bare essentials, we make the simplifying assumption that the op amp has just one single pole, so $\phi_{x(\text{HOR})} = 0$.

The Differentiator

The differentiator of Fig. 8.36*a* is a notorious example of a circuit with a feedback pole. Stemming from the low-pass action provided by the *R*-*C* network to the feedback signal, this pole produces phase lag that adds to that produced by the op amp, reducing ϕ_m to a perilously low value. For a closer look, refer to the PSpice circuit of Fig. 8.37*a*, which uses an op amp with GBP = 1 MHz to provide differentiation with

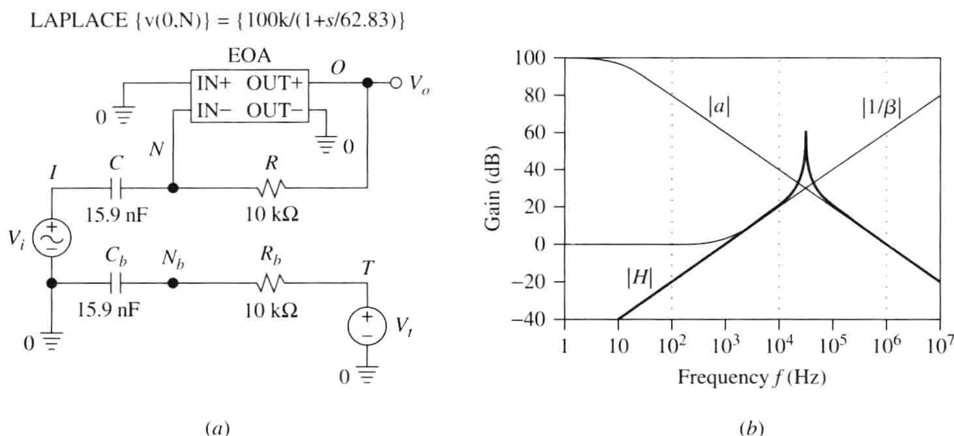


FIGURE 8.37
(*a*) Differentiator designed for a unity-gain frequency of 1.0 kHz, and (*b*) its frequency characteristics. The feedback network for plotting $1/\beta$ is repeated at the bottom of (*a*), with its elements identified by subscript *b*. The $|H|$ trace is plotted as `DB(V(O)/V(I))`, the $|a|$ trace as `DB(V(OA)/(-V(N)))`, and the $|1/\beta|$ trace as `DB(V(T)/V(Nb))`.

unity-gain frequency $f_0 = 1/(2\pi RC) = 1$ kHz. Its response $|H(jf)| = |V_o/V_i|$ is shown in Fig. 8.37*b*, along with the $|a|$ and $|1/\beta|$ curves, so we can visualize the loop gain $|T|$ as the difference between the two. So long as $|T| \gg 1$, we have $H \to H_{\text{ideal}} = -jf/f_0$. However, as we approach the crossover frequency f_x, $|H|$ exhibits considerable peaking, and past f_x, where $|T| \ll 1$, it rolls off with $|a|$. Although $H(jf)$ can be derived analytically (see Problem 8.42), visual inspection reveals that the rate of closure (ROC) approaches 40 dB/dec, so ϕ_m approaches zero (this is similar to the $|1/\beta_2|$ curve of Fig. 8.10). Cursor measurements give $f_x = 31.6$ kHz and $\sphericalangle T(jf_x) = -178.2°$, so $\phi_m = 1.8°$, indicating a circuit on the verge of oscillation.

A popular way of stabilizing the differentiator is by inserting a series resistance R_c as in Fig. 8.36*b*. At low frequencies R_c has little effect because $R_c \ll |Z_C|$. But at high frequencies, where C acts as a short compared to R_c, the noise gain tends to $1/\beta_\infty = 1 + R/R_c$, indicating that the $|1/\beta|$ curve must have a pole frequency. Adjusting R_c so as to place this pole right at the crossover frequency will reduce the ROC and yield $\phi_m \cong 45°$ (this is similar to the $|1/\beta_3|$ curve of Fig. 8.10).

EXAMPLE 8.9. Find R_c to stabilize the differentiator of Fig. 8.37*a* for $\phi_m \cong 45°$. Verify with PSpice.

Solution. Inspection of Fig. 8.37*b* reveals that $|a(jf_x)| = 30$ dB $= 31.6$ V/V. To place the pole right at f_x, we must impose $1/\beta_\infty = 1 + 10^4/R_c = 31.6$, or $R_c = 326\ \Omega$. Running the PSpice circuit of Fig. 8.38*a*, we get the curves of Fig. 8.38*b*, where cursor measurements give $\phi_m = 53.8°$ (even better than the target of 45°).

Stray Input Capacitance Compensation

All real-life op amps exhibit stray input capacitances. Of special concern is the net capacitance C_n of the inverting input toward ground,

$$C_n = C_d + C_c/2 + C_{\text{ext}} \tag{8.31}$$

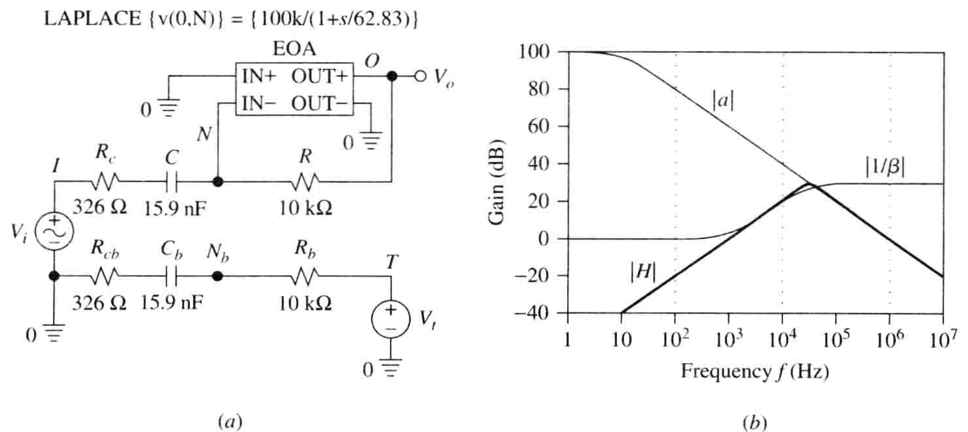


FIGURE 8.38
(*a*) Differentiator compensated for $\phi_m \cong 45°$, and (*b*) its frequency characteristics.

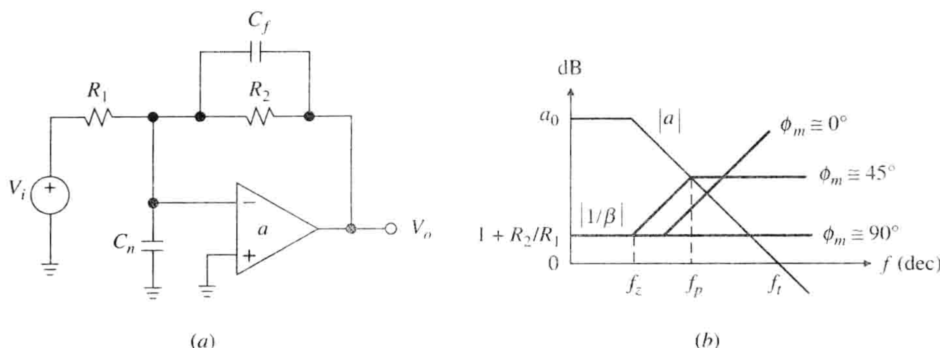


FIGURE 8.39
Using the feedback capacitance C_f to compensate for the stray input capacitance C_n.

where C_d is the *differential capacitance* between the input pins; $C_c/2$ is the *common-mode capacitance* of each input to ground, so that when the inputs are tied together the net capacitance is the sum of the two; and C_{ext} is the *external parasitic capacitance* of components, leads, sockets, and printed-circuit traces associated with the inverting input node. Typically, each of the above components is on the order of a few picofarads.

As in the case of the differentiator, C_n creates a feedback pole whose phase lag erodes ϕ_m. A common way of counteracting this lag is by using a feedback capacitance C_f to create feedback phase lead. This is illustrated in Fig. 8.39*a* for the inverting case. Assuming $r_i = \infty$ and $r_o = 0$, one can easily prove (see Problem 8.43) that the noise gain is

$$\frac{1}{\beta(jf)} = \left(1 + \frac{R_2}{R_1}\right)\frac{1 + jf/f_z}{1 + jf/f_p}, \; f_z = \frac{1}{2\pi(R_1 \| R_2)(C_n + C_f)}, \; f_p = \frac{1}{2\pi R_2 C_f} \tag{8.32}$$

In the absence of C_f, the noise gain exhibits a breakpoint at f_z because of C_n. If f_z is sufficiently low to make the ROC approach 40 dB/dec, the circuit will exhibit excessive peaking and ringing. Physically, this is not surprising as the impedance presented by the inverting input is of the shunt type and thus inductive, so C_n will tend to resonate with the inductive component L_{eq} (see Fig. 6.11*b*).

Inserting C_f lowers f_z somewhat while also establishing a second breakpoint at f_p, above which the noise gain flattens out toward the high-frequency asymptote $1/\beta_\infty = 1 + Z_{C_f}/Z_{C_n} = 1 + C_n/C_f$. By properly positioning this second breakpoint, we can increase ϕ_m. For $\phi_m \cong 45°$ we place f_p right on the $|a|$ curve, so $f_p = \beta_\infty f_t$. Rewriting as $1/(2\pi R_2 C_f) = f_t/(1 + C_n/C_f)$ gives

$$C_f = (1 + \sqrt{1 + 8\pi R_2 C_n f_t})/(4\pi R_2 f_t) \qquad \text{for } \phi_m \cong 45° \tag{8.33a}$$

Alternatively, we can compensate for $\phi_m = 90°$. In this case we place f_p right on top of f_z so as to cause a pole-zero cancellation. This makes the $|1/\beta|$ curve flat throughout, or $1/\beta_\infty = 1 + R_2/R_1$. Rewriting as $1 + C_n/C_f = 1 + R_2/R_1$ yields

$$C_f = (R_1/R_2)C_n \qquad \text{for } \phi_m = 90° \tag{8.33b}$$

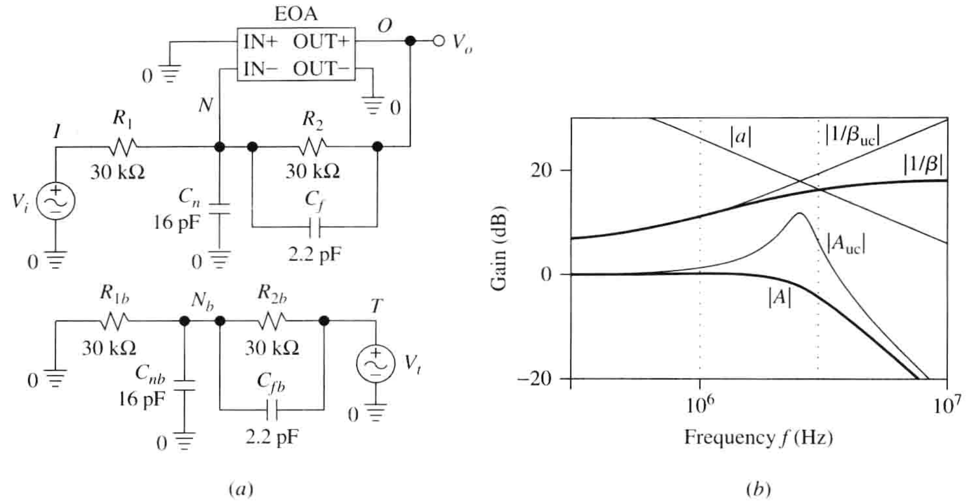


FIGURE 8.40
(*a*) Stray input capacitance compensation, and (*b*) frequency characteristics.

This alternative, also referred to as *neutral compensation*, is similar to the compensation of oscilloscope probes. However, the price for $\phi_m \cong 90^\circ$ is a lower closed-loop bandwidth of f_p.

EXAMPLE 8.10. Let the op amp of Fig. 8.39*a* have $a_0 = 10^5$ V/V, $f_t = 20$ MHz, $C_d = 7$ pF, and $C_c/2 = 6$ pF. (*a*) If $R_1 = R_2 = 30$ kΩ and $C_{\text{ext}} = 3$ pF, use PSpice to verify that the circuit does not have an adequate phase margin. (*b*) Find C_f for $\phi_m \cong 45^\circ$, calculate f_B, and verify with PSpice. (*c*) Repeat (*b*), but for the case of neutral compensation. Compare with (*b*).

Solution.

(*a*) We have $C_n = 7 + 6 + 3 = 16$ pF. Running the PSpice circuit of Fig. 8.40*a* with $C_f = 0$ we get the uncompensated curves $|A_{\text{uc}}|$ and $|1/\beta_{\text{uc}}|$ of Fig. 8.40*b*. Their crossover frequency is 2.53 MHz, where the combined phase shift by a and β_{uc} is -165.3°, so $\phi_m = 14.7^\circ$, a poor phase margin.

(*b*) Using Eq. 8.33*a*, we find $C_f = 2.2$ pF. Moreover, $f_B = 1/(2\pi \times 30 \times 10^3 \times 2.2 \times 10^{-12}) = 2.4$ MHz. Rerunning PSpice with C_f in place, we get the $|A|$ and $|1/\beta|$ curves, whose crossover frequency is 3.03 MHz. The corresponding phase shift is -117.6°, so $\phi_m = 62.4^\circ$ (even better than the target of 45°). Moreover, PSpice gives $f_B = 2.7$ MHz.

(*c*) Use $C_f = 16$ pF and get $f_B = 0.332$ MHz (PSpice gives $f_B = 0.330$ MHz), a much lower bandwidth.

We now turn to the noninverting configuration[10] of Fig. 8.41*a*, where the various stray input capacitances have been shown explicitly. We observe that the overall capacitance C_n is still given by Eq. (8.31). However, the portion $C_1 = C_c/2 + C_{\text{ext}}$ is now in parallel with R_1, so we have $A_{\text{ideal}} = 1 + Z_2/Z_1$, $Z_1 = R_1 \parallel [1/(j2\pi f C_1)]$,

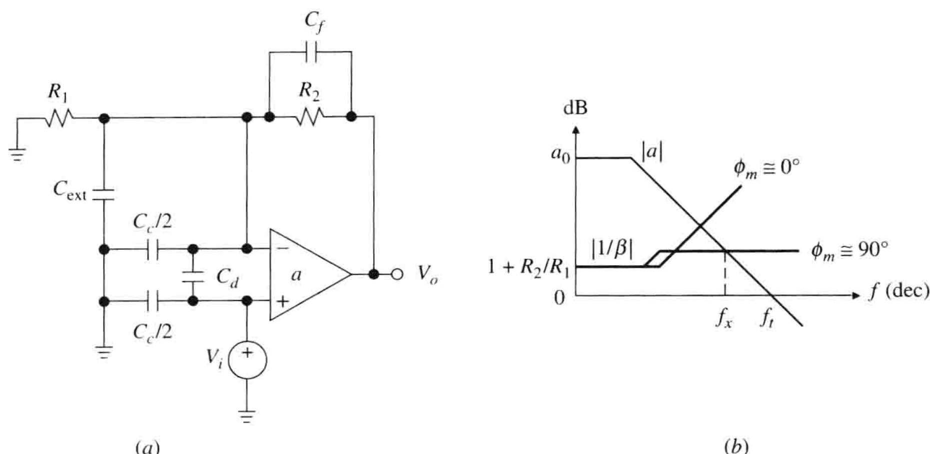


FIGURE 8.41
Stray input capacitance compensation for the noninverting configuration.

$Z_2 = R_2 \parallel [1/(j2\pi f C_f)]$. We can make A_{ideal} frequency-independent by using

$$C_f = (R_1/R_2)(C_c/2 + C_{\text{ext}}) \tag{8.34}$$

The effect of C_f is shown in Fig. 8.41*b*. The actual gain is now $A(jf) \cong (1 + R_2/R_1)/(1 + jf/f_x)$, $f_x = \beta_\infty f_t = f_t/(1 + C_n/C_f)$.

EXAMPLE 8.11. Stabilize the circuit of Fig. 8.41*a* if the data are the same as in Example 8.10. Hence, find $A(jf)$.

Solution. We have $C_f = (30/30)(6+3) = 9\,\text{pF}$, $f_x = 2\times 10^7/(1+16/9) = 7.2\,\text{MHz}$, and

$$A(jf) \cong \frac{2}{1 + jf/(7.2\,\text{MHz})}\ \text{V/V}$$

With careful component layout and wiring, C_{ext} can be minimized but not altogether eliminated. Consequently, it is always a good practice to include a small feedback capacitance C_f in the range of a few picofarads to combat the effect of C_n as given in Eq. (8.31).

Capacitive-Load Isolation

There are applications in which the external load is heavily capacitive. Sample-and-hold amplifiers and peak detectors are typical examples. When an op amp drives a coaxial cable, it is the distributed cable capacitance that makes the load capacitive. To investigate capacitive loading, consider the simple circuit of Fig. 8.42*a*. As shown more explicitly in Fig. 8.42*b* (ignore the C_s-R_s network for a moment), the load C_L forms a pole with the op amp's output resistance z_o, thus increasing the phase lag around the loop. This reduces the phase margin, possibly causing peaking and ringing. Again, we justify this physically by noting that the output impedance presented by the follower is of the shunt type and thus inductive, so C_L will tend

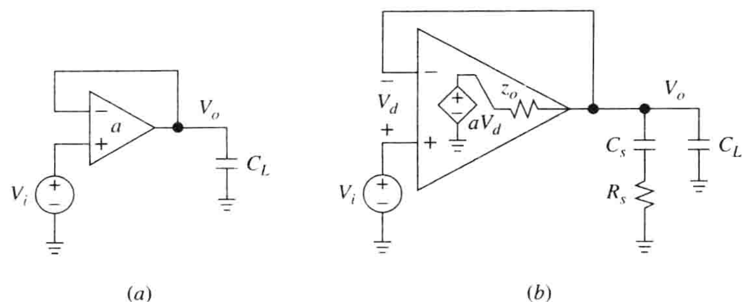


FIGURE 8.42
(*a*) Capacitively loaded voltage follower, and (*b*) stabilization using a snubber network.

to resonate with the inductive component L_{eq} (see Fig. 6.11*b*). To tame the ringing, we must alter the damping conditions, a function that is provided by the C_s-R_s snubber.

As an example, consider the PSpice circuit of Fig. 8.43, which uses an op amp with $f_t = 1$ MHz and $z_o = 100\ \Omega$, and is loaded by $C_L = 50$ nF. Since $\beta = 1$, $T = a$ in this case. Without the snubber, the circuit gives the uncompensated (uc) responses of Fig. 8.44. The loop gain T_{uc} contains two pole frequencies, one at 10 Hz and the other at $1/(2\pi z_o C_L) \cong 32$ kHz. Its ROC with the 0-dB line approaches 40 dB/dec, indicating an almost oscillatory circuit. Cursor measurements give $f_x = 177$ kHz and $\measuredangle T_{uc}(jf_x) = -169.8^\circ$, so $\phi_m = 10.2^\circ$, which explains the pronounced peaking of $|A_{uc}|$ and ringing of $v_{O(uc)}$.

The open-loop output impedance z_o is small and resistive at dc, but tends to become a complex function of frequency. Consequently, the snubber network values must be found empirically. Starting out with a large C_s, we adjust R_s until the gain peaking (GP) and the overshoot (OS) are lowered to their desired levels. Then, we lower C_s to a value that still retains acceptable GP/OS characteristics. With the snubber values shown in Fig. 8.43, the ROC of the $|T|$ curve with the 0-dB line is appreciably lower. In fact, cursor measurements now give $f_x = 115$ kHz and $\measuredangle T(jf_x) = -119^\circ$, so $\phi_m = 61^\circ$, which explains the much reduced peaking of $|A|$ and ringing of v_O.

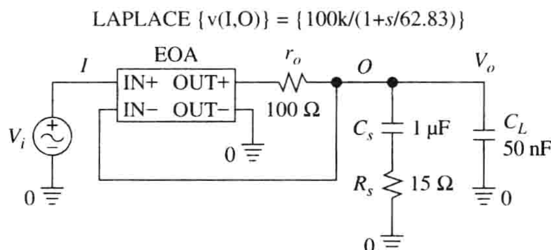


FIGURE 8.43
PSpice circuit to investigate the capacitive loading of a voltage follower.

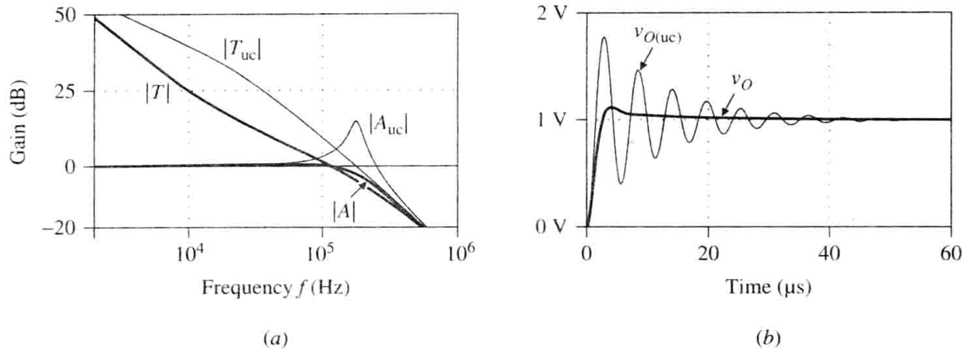


FIGURE 8.44
(*a*) Frequency characteristics and (*b*) step responses of the follower of Fig. 8.43.

Figure 8.45 shows a capacitive-load stabilization alternative that is popular with resistive circuits such as the inverting/noninverting amplifiers and the summing/difference amplifiers. Also called *in-loop compensation*, it uses a small series resistance R_s to decouple the op amp's output node from C_L, and a suitable feedback capacitance C_f to provide a high-frequency bypass around C_L as well as to combat the effect of any stray input capacitance C_n. It is possible to specify the compensation network so that the phase lead introduced by C_f exactly neutralizes the phase lag due to C_L. The design equations for neutral compensation are[11]

$$R_s = (R_1/R_2)r_o \qquad C_f = (1 + R_1/R_2)^2(r_o/R_2)C_L \tag{8.35a}$$

Moreover, the closed-loop bandwidth is[12]

$$f_B = \frac{1}{2\pi(1 + R_2/R_1)R_sC_L} \tag{8.35b}$$

and is set by the external components independently of the op amp's GBP.

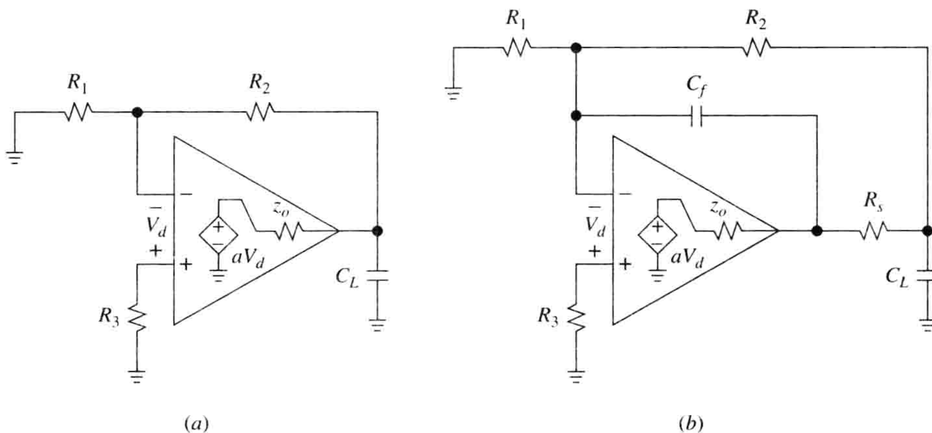


FIGURE 8.45
(*a*) Generalized resistive-feedback circuit, and (*b*) in-loop compensation.

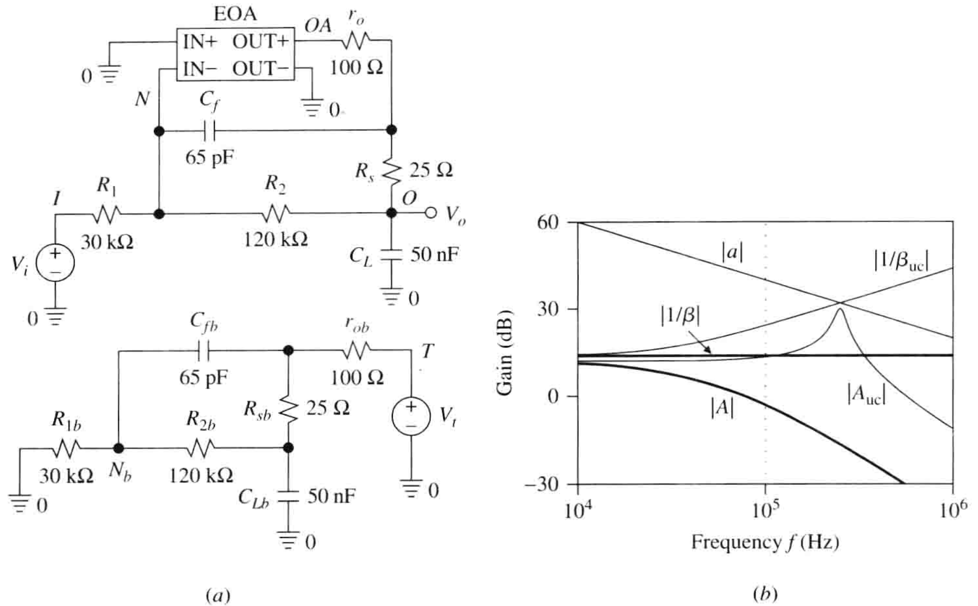


FIGURE 8.46
(*a*) PSpice circuit of Example 8.12, and (*b*) its frequency characteristics.

EXAMPLE 8.12. Suppose the circuit of Fig. 8.45*a* is configured as an inverting-amplifier with $R_1 = 30\,\text{k}\Omega$ and $R_2 = 120\,\text{k}\Omega$, and it drives a 50-nF load. (*a*) If the op amp has $a_0 = 10^5$ V/V, $f_t = 10$ MHz, and $r_o = 100\,\Omega$, find R_s and C_f for neutral compensation. What is the closed-loop bandwidth? (*b*) Verify with PSpice, compare with the uncompensated case, and comment.

Solution.

(*a*) By Eq. (8.35), $R_s = (30/120)100 = 25\,\Omega$, $C_f = (1 + 30/120)^2(0.1/120)50 \times 10^{-9} = 65$ pF, $f_B = 1/[2\pi(1 + 120/30)25 \times 50 \times 10^{-9}] = 25.5$ kHz.

(*b*) Running the PSpice circuit of Fig. 8.46*a*, we get the plots of Fig. 8.46*b*. Without compensation the circuit has $\phi_m = 7.2°$, with compensation it has $\phi_m \cong 90°$ and $f_B = 24$ kHz. Even though the crossover frequency is above 1 MHz, f_B is set by the external components and is much lower, this being the price we are paying for neutral compensation.

As mentioned, at high frequencies z_o is no longer purely resistive, so the above equations offer only initial estimates for R_s and C_f. The optimum values must be found empirically once the circuit has been assembled in the lab.

The need to drive capacitive loads arises frequently enough to warrant the design of special op amps with provisions for automatic capacitive-load compensation. The AD817 and LT1360 op amps are designed to drive unlimited capacitive loads. Special internal circuitry senses the amount of loading and adjusts the open-loop response to maintain an adequate phase margin regardless of the load. The process, completely transparent to the user, is most effective when the load is not fixed or is ill-defined, as in the case of unterminated coaxial cable loads.

Other Sources of Instability

In circuitry incorporating high-gain amplifiers such as op amps and voltage comparators, the specter of instability arises in a number of subtle ways unless proper circuit design and construction rules are followed.[13-16] Two common causes of instability are *poor grounding* and *inadequate power-supply filtering*. Both problems stem from the distributed impedances of the supply and ground busses, which can provide spurious feedback paths around the high-gain device and compromise its stability.

In general, to minimize the ground-bus impedance, it is good practice to use a ground plane, especially in audio and wideband applications. To reduce grounding problems further, it is good practice to provide two separate ground busses: a *signal-ground* bus to provide a return path for critical circuits—such as signal sources, feedback networks, and precision voltage references—and a *power-ground* bus to provide a return path for less critical circuits, such as high-current loads and digital circuits. Every effort is made to keep both dc and ac currents on the signal-ground bus *small* in order to render this bus essentially equipotential. To avoid perturbing this equipotential condition, the two busses are joined only at one point of the circuit.

Spurious feedback paths can also form through the power-supply busses. Because of nonzero bus impedances, any change in supply currents brought about by a load current change will induce a corresponding voltage change across the op amp supply pins. Due to finite PSRR, this change will in turn be felt at the input, thus providing an indirect feedback path. To break this path, each supply voltage must be bypassed with a 0.01-μF to 0.1-μF decoupling capacitor, in the manner already depicted in Fig. 1.43. The best results are obtained with low ESR and ESL ceramic chip capacitors, preferably surface-mounted. For this cure to be effective, the lead lengths must be kept short and the capacitors must be mounted as close as possible to the op amp pins. Likewise, the elements of the feedback network must be mounted close to the inverting-input pin in order to minimize the stray capacitance C_{ext} appearing in Eq. (8.31). Manufacturers often provide evaluation boards to guide the user in the proper construction of the circuit.

8.5 INPUT-LAG AND FEEDBACK-LEAD COMPENSATION

Input-lag and feedback-lead are popular compensation techniques that are used to stabilize a circuit by lowering its ROC via suitable manipulation of the noise gain $1/\beta(jf)$.

Input-Lag Compensation

This technique, depicted in Fig. 8.47*a* for a generalized resistive-feedback circuit, reshapes the noise gain by connecting a suitable C_c-R_c network across the op amp's input terminals. At low frequencies C_c acts as an open circuit, so the compensation network has no effect, yielding the familiar low-frequency noise gain $1/\beta_0 = 1 + R_2/R_1$. However, at high frequencies, where C_c acts as a short, the noise gain shoots up to the asymptotic value (see Problem 8.52)

$$\frac{1}{\beta_\infty} = 1 + \frac{R_2}{R_1} + \frac{R_2 + (1 + R_2/R_1)R_3}{R_c}$$

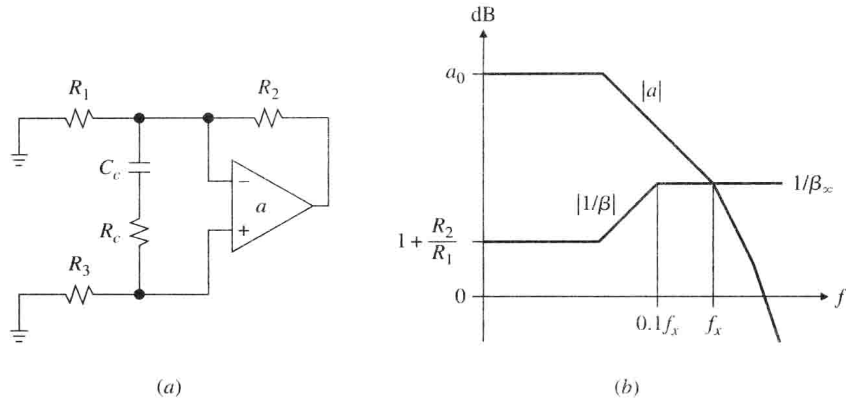


FIGURE 8.47
(*a*) Input-lag compensation, (*b*) illustration via Bode plots for the case $\phi_m \cong 45°$.

Once we decide on the value of f_x based on the desired noise margin, we adjust R_c so as to make $1/\beta_\infty = |a(jf_x)|$, as exemplified in Fig. 8.47*b* for the case $\phi_m \cong 45°$. The design equation for R_c is thus

$$R_c = \frac{R_2 + (1 + R_2/R_1)R_3}{|a(jf_x)| - (1 + R_2/R_1)} \tag{8.36a}$$

The high-frequency breakpoint of the noise margin (see Problem 8.52) is $1/(2\pi R_c C_c)$, this being the frequency at which the impedance of C_c equals, in magnitude, R_c. To contain the erosion of ϕ_m, it is customary to position this frequency about a decade below f_x, as shown. Consequently,

$$C_c = \frac{5}{\pi R_c f_x} \tag{8.36b}$$

The input-lag technique is used to stabilize uncompensated as well as decompensated op amps, and also circuits that use compensated op amps but have feedback poles, like capacitively loaded circuits.

EXAMPLE 8.13. A difference amplifier is implemented with four 10-kΩ resistors and an uncompensated op amp having $a_0 = 10^5$ V/V and two pole frequencies of 1 kHz and 1 MHz. (*a*) Verify that the circuit needs compensation. (*b*) Design an input-lag network to stabilize it for $\phi_m = 60°$, and verify with PSpice.

Solution.

(*a*) With a noise gain as low as $1 + 10/10 = 2$ V/V, the crossover frequency of the uncompensated (uc) circuit is in the region where the ROC approaches 40-dB/dec, indicating a negligible phase margin. Indeed, using the PSpice circuit of Fig. 8.48*a* but without R_c and C_c, we find $\phi_m \cong 8°$ (see Fig. 8.48*b*).

(*b*) For $\phi_m = 60°$ we need $f_x = f_{-120°}$. Using again the cursor, we find $f_{-120°} =$ 581 kHz, and $|a(jf_{-120°})| = 43.5$ dB, or 149 V/V. Substituting these values into Eqs. (8.36), along with $R_3 = 10 \parallel 10 = 5$ k Ω, we get $R_c = 136\ \Omega$ and $C_c = 20$ nF. Rerunning PSpice with R_c and C_c in place, we get $f_x = 584$ kHz and $\measuredangle T(jf_x) = -125.7°$, so $\phi_m = 54.3°$.

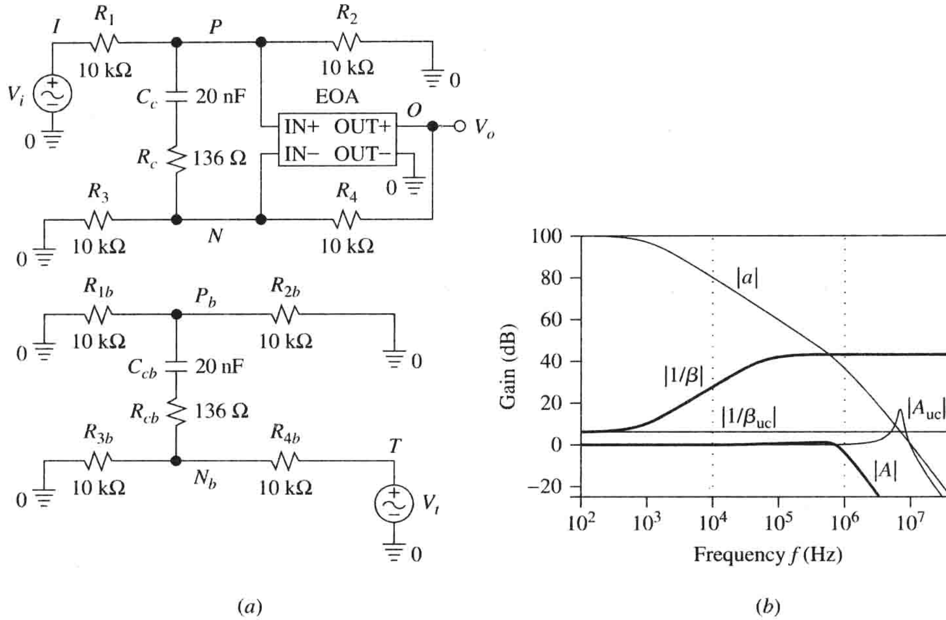


FIGURE 8.48
(a) Input-lag compensated difference amplifier of Example 8.13, and (b) its frequency characteristics. The $|A|$ trace is plotted as `DB(V(O)/V(I))`, the $|a|$ trace as `DB(V(O)/(V(P)-V(N)))`, and the $|1/\beta|$ trace as `DB(V(T)/(V(Nb)-V(Pb)))`.

A popular application of input-lag compensation is the stabilization of decompensated op amps for operation all the way to $\beta = 1$. A classic example is the LF356/357 op amp pair: the 356 version uses an on-chip capacitor $C_c = 10$ pF to provide GBP = 5 MHz and SR = 12 V/μs for $\beta \leq 1$; the 357, the decompensated version, uses $C_c = 3$ pF to achieve faster dynamics (GBP = 20 MHz and SR = 50 V/μs) but only for operation with $\beta \leq 0.2$, or noise gains of $1/\beta \geq 5$ V/V (5.5 V/V = 14 dB). What if we want to operate a 357 with $1/\beta \leq 5$ V/V? Figure 8.49a shows

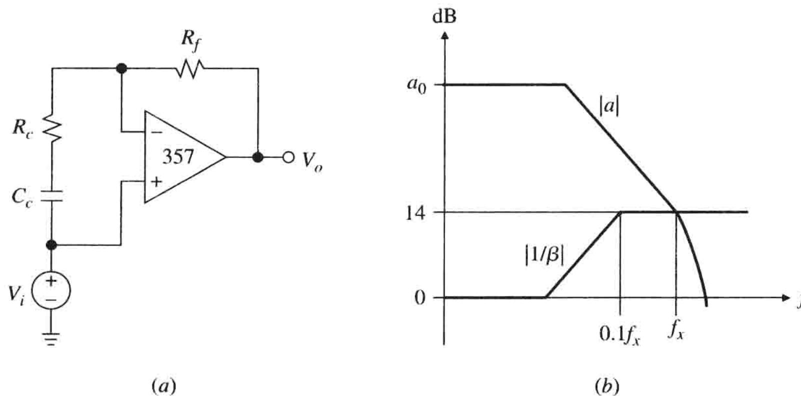


FIGURE 8.49
(a) Using a decompensated op amp as a voltage follower, and (b) Bode plots.

the use of input-lag compensation to stabilize a 357 for voltage-follower ($\beta = 1$) operation. Adapting Eq. (8.36*a*) to the present case, we get $R_c = R_f/(5-1) = R_f/4$ (use $R_f = 12$ kΩ and $R_c = 3$ kΩ). Then, adapting Eq. (8.36*b*) with $f_x = \text{GBP}/5 = 20/5 = 4$ MHz, we get $C_c = 5/(\pi \times 3 \times 10^3 \times 4 \times 10^6) = 133$ pF (use 130 pF).

Input-lag compensation provides yet another alternative for the stabilization of capacitively loaded circuits, as demonstrated by the following example.

EXAMPLE 8.14. Use input-lag compensation to stabilize the circuit of Example 8.12 for $\phi_m = 45°$.

Solution. We use the PSpice circuit of Fig. 8.50*a*, but without the R_c-C_c network, to plot $|a_{\text{loaded}}|$, the gain of the op amp loaded by C_L. Next, we use the cursor to find $f_{-135°} = 32$ kHz and $|a_{\text{loaded}}(jf_{-135°})| = 46.836$ dB, or 220 V/V. Proceeding as in Example 8.13, we get $R_c = 559\ \Omega$ and $C_c = 89$ nF. Using again the cursor, we find $\phi_m = 39.4°$ (the 5.6° phase-margin erosion from the target value of 45° is due to the presence of the breakpoint at $0.1 f_x$).

Compared with internal compensation, the input-lag method allows for higher slew rates as the op amp is spared from having to charge or discharge any internal compensating capacitance. The capacitance is now connected between the inputs, so the voltage changes it experiences tend to be very small. However, the settling-time improvement stemming from a higher slew rate is counterbalanced by a long settling tail[17] due to the presence of a pole-zero doublet at f_z and f_p.

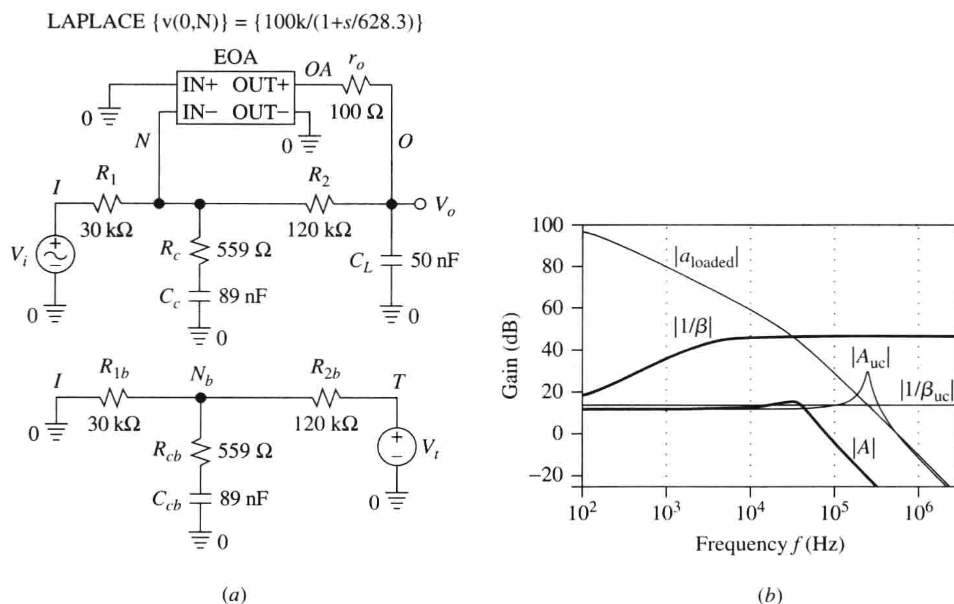


FIGURE 8.50
(*a*) Input-lag compensation of the capacitively loaded circuit of Example 8.12, and (*b*) its frequency characteristics. The $|a_{\text{loaded}}|$ trace is plotted as `DB(V(O)/(-V(N)))`, and the $|1/\beta|$ trace as `DB(V(T)/V(Nb))`.

A notorious disadvantage of this method is increased high-frequency noise, since the noise-gain curve is raised significantly in the vicinity of the crossover frequency f_x. Another disadvantage is a much lower closed-loop differential input impedance Z_d, since z_d is now in parallel with $Z_c = R_c + 1/(j2\pi f C_c)$, and Z_c is much smaller than z_d. Though this is inconsequential in inverting configurations, it may cause intolerable high-frequency input loading and feedthrough in noninverting configurations. Input-lag compensation is nevertheless popular.

Feedback-Lead Compensation

This technique[1] uses a feedback capacitance C_f to create phase lead in the feedback path. This lead is designed to occur in the vicinity of the crossover frequency f_x, which is where ϕ_m needs to be boosted. Alternatively, we can view this method as a reshaping of the $|1/\beta|$ curve near f_x to reduce the rate of closure ROC. Referring to Fig. 8.51*a* and assuming $r_d = \infty$ and $r_o = 0$, we have $1/\beta = 1 + Z_2/R_1$, $Z_2 = R_2 \parallel [1/(j2\pi f C_f)]$. Expanding, we can write

$$\frac{1}{\beta(jf)} = \left(1 + \frac{R_2}{R_1}\right)\frac{1 + jf/f_z}{1 + jf/f_p} \tag{8.37}$$

where $f_p = 1/(2\pi R_2 C_f)$ and $f_z = (1+R_2/R_1)f_p$. As depicted in Fig. 8.51*b*, $|1/\beta|$ has the low- and high-frequency asymptotes $|1/\beta_0| = 1 + R_2/R_1$ and $|1/\beta_\infty| = 0$ dB, and two breakpoints at f_p and f_z.

The phase lag provided by $1/\beta(jf)$ is maximum[1] at the geometric mean of f_p and f_z, so the optimum value of C_f is the one that makes this mean coincide with the crossover frequency, or $f_x = \sqrt{f_p f_z} = f_p\sqrt{1 + R_2/R_1}$. Under such a condition we have $|a(jf_x)| = \sqrt{1 + R_2/R_1}$, which can be used to find f_x via trial and error. Once f_x is known, we find $C_f = 1/(2\pi R_2 f_p)$, or

$$C_f = \frac{\sqrt{1 + R_2/R_1}}{2\pi R_2 f_x} \tag{8.38}$$

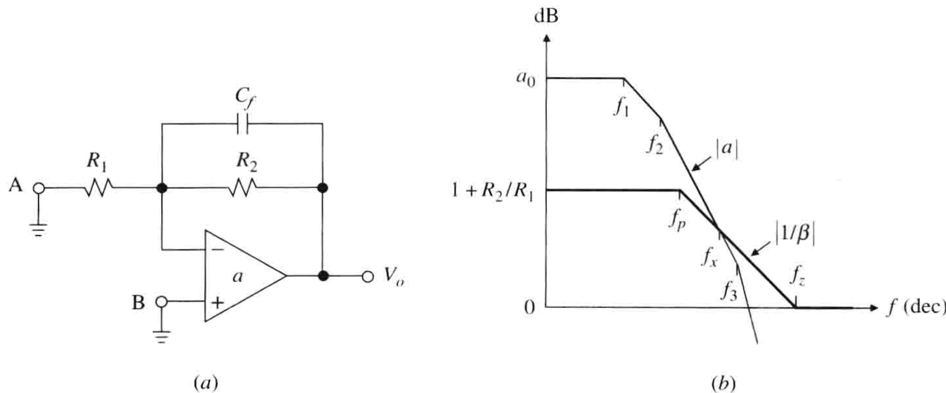


FIGURE 8.51
Feedback-lead compensation.

The closed-loop bandwidth is approximately $1/(2\pi R_2 C_f)$. Moreover, C_f helps combat the effect of the inverting-input stray capacitance C_n.

One can readily verify that at the geometric mean of f_p and f_z we have $\measuredangle(1/\beta) = 90° - 2\tan^{-1}\sqrt{1 + R_2/R_1}$, so the larger the value of $1 + R_2/R_1$, the greater the contribution of $1/\beta$ to ϕ_m. For example, with $1 + R_2/R_1 = 10$ we get $\measuredangle(1/\beta) = 90° - 2\tan^{-1}\sqrt{10} \cong -55°$, which yields $\measuredangle T = \measuredangle a - (-55°) = \measuredangle a + 55°$. We observe that for this compensation scheme to work for a given ϕ_m, the open-loop gain must satisfy $\measuredangle a(jf_x) \geq \phi_m - 90° - 2\tan^{-1}\sqrt{1 + R_2/R_1}$.

EXAMPLE 8.15. (*a*) Using an op amp with $a_0 = 10^5$ V/V, $f_1 = 1$ kHz, $f_2 = 100$ kHz, and $f_3 = 5$ MHz, design a noninverting amplifier with $A_0 = 20$ V/V. Hence, verify that the circuit needs compensation. (*b*) Stabilize it with the feedback-lead method, and find ϕ_m. (*c*) Find the closed-loop bandwidth.

Solution.

(*a*) For $A_0 = 20$ V/V use $R_1 = 1.05\ \text{k}\Omega$ and $R_2 = 20.0\ \text{k}\Omega$. Then $\beta_0 = 1/20$ V/V, and $a_0\beta_0 = 10^5/20 = 5000$. Thus, without compensation we have

$$T(jf) = \frac{5000}{[1 + jf/10^3][1 + jf/10^5][1 + jf/(5 \times 10^6)]}$$

Using trial and error, we find that $|T| = 1$ for $f = 700$ kHz, and that $\measuredangle T(j700\ \text{kHz}) = -179.8°$. So, $\phi_m = 0.2°$, indicating a circuit in bad need of compensation.

(*b*) Again using trial and error, we find that $|a| = \sqrt{20}$ V/V for $f = 1.46$ MHz, and $\measuredangle a(j1.46\ \text{MHz}) = -192.3°$. Letting $f_x = 1.46$ MHz in Eq. (8.38) yields $C_f = 24.3$ pF. Moreover, $\phi_m = 180° + \measuredangle a - (90° - 2\tan^{-1}\sqrt{20}) = 180° + (-192.3°) - (90° - 2 \times 77.4°) = 52.5°$.

(*c*) $f_{-3\text{dB}} = 1/(2\pi R_2 C_f) = 327$ kHz.

We observe that feedback-lead compensation does not enjoy the slew-rate advantages of input-lag compensation; however, it provides better filtering capabilities for internally generated noise. These are some of the factors the user needs to consider when deciding which method is best for a given application.

8.6 STABILITY IN CFA CIRCUITS[18]

The open-loop response $z(jf)$ of a current-feedback amplifier (CFA) is dominated by a single pole only over a designated frequency band. Beyond this band, higher-order roots come into play, which increase the overall phase shift. When frequency-independent feedback is applied around a CFA, the latter will offer unconditional stability with a specified phase margin ϕ_m only as long as $1/\beta \geq (1/\beta)_{\min} = |z(jf_{\phi_m - 180°})|$, where $f_{\phi_m - 180°}$ is the frequency at which $\measuredangle z = \phi_m - 180°$. Lowering the $1/\beta$ curve below $(1/\beta)_{\min}$ would increase the phase shift, thus eroding ϕ_m and inviting instability. This behavior is similar to that of decompensated op amps. The value of $(1/\beta)_{\min}$ can be found from the data-sheet plots of $|z(jf)|$ and $\measuredangle z(jf)$. As with voltage-feedback amplifiers (VFAs), instability in CFA circuits may also stem from feedback phase lag due to external reactive elements.

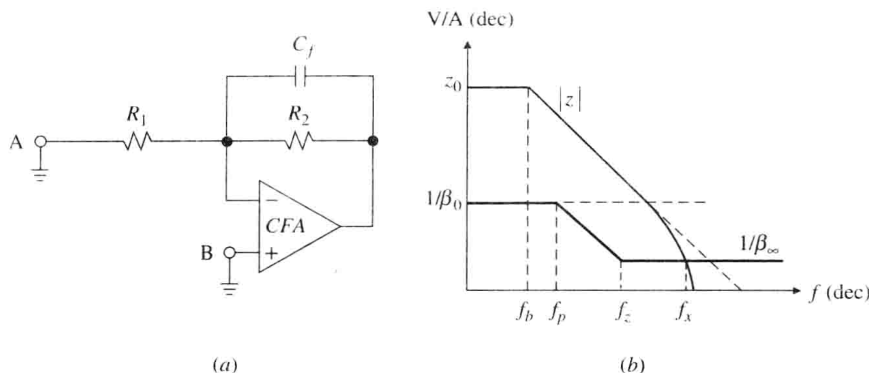


FIGURE 8.52
A large feedback capacitance C_f tends to destabilize a CFA circuit.

Effect of Feedback Capacitance

To investigate the effect of feedback capacitance, refer to Fig. 8.52*a*. At low frequencies, C_f acts as an open circuit, so we can apply Eq. (6.64) and write $1/\beta_0 = R_2 + r_n(1 + R_2/R_1)$. At high frequencies, R_2 is shorted out by C_f, so $1/\beta_\infty = 1/\beta_0|_{R_2\to 0} = r_n$. Since $1/\beta_\infty \ll 1/\beta_0$, the crossover frequency f_x is pushed into the region of greater phase shift, as shown in Fig. 8.52*b*. If this shift reaches $-180°$, the circuit will oscillate.

We thus conclude that *direct capacitive feedback must be avoided in CFA circuits.* In particular, the familiar inverting or Miller integrator is not amenable to CFA implementation, unless suitable measures are taken to stabilize it (see Problem 8.62). However, the noninverting or Deboo integrator is acceptable because β in the vicinity of f_x is still controlled by the resistance in the negative-feedback path. Likewise, we can readily use CFAs in those filter configurations that do not employ any direct capacitance between the output and the inverting input, such as *KRC* filters.

Stray Input Capacitance Compensation

In Fig. 8.53*a* C_n appears in parallel with R_1. Replacing R_1 with $R_1 \parallel [1/(j2\pi f C_n)]$ in Eq. (6.64) yields, after minor algebra,

$$\frac{1}{\beta} = \frac{1}{\beta_0}(1 + jf/f_z) \tag{8.39a}$$

$$\frac{1}{\beta_0} = R_2 + r_n\left(1 + \frac{R_2}{R_1}\right) \qquad f_z = \frac{1}{2\pi(R_1 \parallel R_2 \parallel r_n)C_n} \tag{8.39b}$$

As shown in Fig. 8.53*b*, the $1/\beta$ curve starts to rise at f_z, and if C_n is sufficiently large to make $f_z < f_x$, the circuit will become unstable.

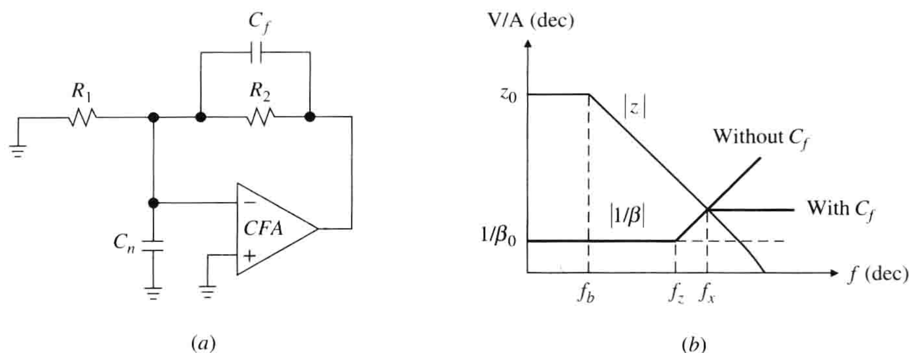


FIGURE 8.53
Input stray capacitance compensation in CFA circuits.

Like a VFA, a CFA is stabilized by counteracting the effect of C_n with a small feedback capacitance C_f. Together with R_2, C_f creates a pole frequency for $1/\beta$ at $f_p = 1/(2\pi R_2 C_f)$. For $\phi_m = 45°$, impose $f_p = f_x$. We observe that f_x is the geometric mean of f_z and $\beta_0 z_0 f_b$. Letting $1/(2\pi R_2 C_f) = \sqrt{\beta_0 z_0 f_b f_z}$ and solving, we get

$$C_f = \sqrt{r_n C_n/(2\pi R_2 z_0 f_b)} \tag{8.40}$$

A typical application is when a CFA is used in conjunction with a current-output DAC to perform fast *I-V* conversion, and the stray capacitance is the combined result of the stray output capacitance of the DAC and the stray inverting-input capacitance of the CFA (see Fig. 8.54).

EXAMPLE 8.16. A current-output DAC is fed to a CFA having $z_0 = 750$ kΩ, $f_b =$ 200 kHz, and $r_n = 50$ Ω. Assuming $R = 1.5$ kΩ and $C_n = 100$ pF, find C_f for $\phi_m = 45°$. Verify with PSpice.

Solution. $C_f = \sqrt{50 \times 100 \times 10^{-12}/(2\pi \times 1.5 \times 10^3 \times 1.5 \times 10^{11})} = 1.88$ pF. This value can be increased for a greater phase margin, but this will also reduce the bandwidth of the *I-V* converter.

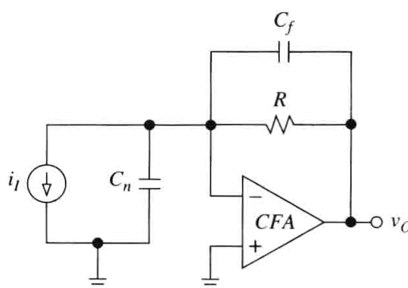


FIGURE 8.54
CFA-based *I-V* converter.

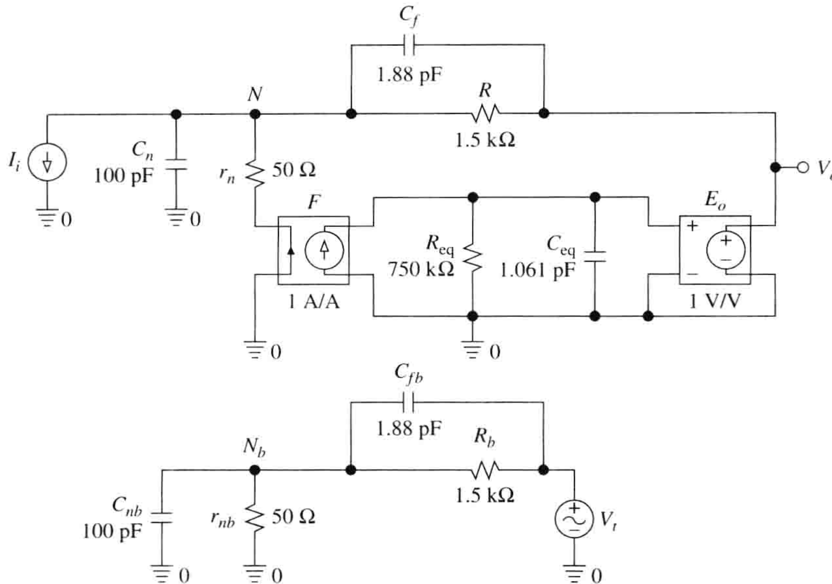


FIGURE 8.55
PSpice circuit for the CFA I-V converter of Example 8.16. Circuit proper (top) and feedback network for plotting $1/\beta$ (bottom).

Using the simplified CFA model of Fig. 6.37, we set up the PSpice circuit of Fig. 8.55 (since the noninverting input is grounded, we can drop the input buffer and short r_n to ground). After running PSpice, we obtain the plots of Fig. 8.56. The uncompensated circuit (uc) has insufficient phase margin ($\phi_m = 32.6°$), which accounts for the peaking and ringing. After compensation we have $\phi_m \cong 74°$, indicating the absence of peaking and negligible ringing. The -3-dB closed-loop frequency is near 70 MHz.

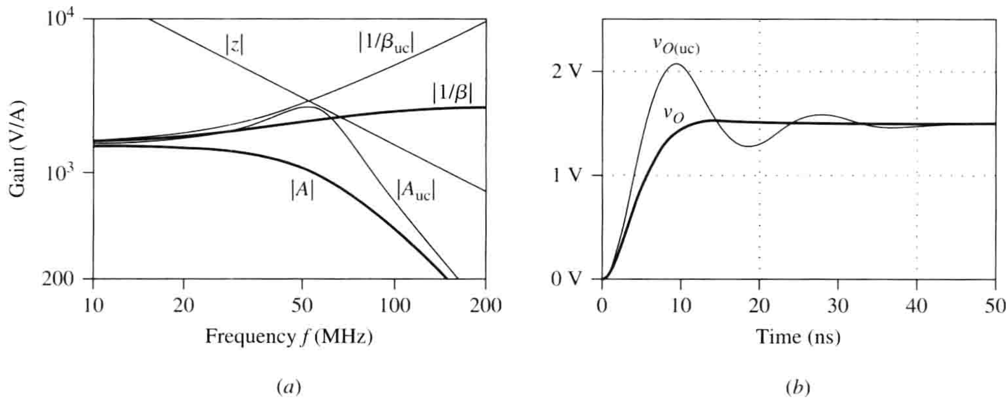


FIGURE 8.56
(a) Frequency and (b) transient responses of the I-V converter of Fig. 8.55. The $|A|$ trace is plotted on a log scale as `V(O)/I(Ii)`, the $|z|$ trace as `V(O)/I(rn)`, and the $|1/\beta|$ trace as `V(T)/(-I(rnb))`.

8.7 COMPOSITE AMPLIFIERS

Two or more op amps can be combined to achieve improved overall performance.[19] The designer need be aware that when an op amp is placed within the feedback loop of another, stability problems may arise. In the following we shall designate the gains of the individual op amps as a_1 and a_2, and the gain of the composite device as a.

Increasing the Loop Gain

Two op amps, usually from a dual-op amp package, can be connected in cascade to create a composite amplifier with a gain $a = a_1 a_2$ much higher than the individual gains a_1 and a_2. We expect the composite device to provide a much greater loop gain, and thus a much lower gain error. However, if we denote the individual unity-gain frequencies as f_{t1} and f_{t2}, we observe that at high frequencies, where $a = a_1 a_2 \cong (f_{t1}/jf)(f_{t2}/jf) = -f_{t1} f_{t2}/f^2$, the phase shift of the composite response approaches $-180°$, thus requiring frequency compensation.

In applications with sufficiently high closed-loop dc gains, the composite amplifier can be stabilized via the feedback-lead method[20] shown in Fig. 8.57*a*. As usual, the circuit can be either an inverting or a noninverting amplifier, depending on whether we insert the input source at node *A* or *B*. The decibel plot of $|a|$ is obtained by adding together the individual decibel plots of $|a_1|$ and $|a_2|$. This is illustrated in Fig. 8.57*b* for the case of matched op amps, or $a_1 = a_2$.

As we know, the $1/\beta$ curve has a pole frequency at $f_p = 1/(2\pi R_2 C_f)$ and a zero frequency at $f_z = (1 + R_2/R_1) f_p$. For ROC = 30 dB/dec, or $\phi_m = 45°$, we place f_p right on the $|a|$ curve. This yields $1 + R_2/R_1 = |a(jf_p)| = f_{t1} f_{t2}/f_p^2$. Solving for f_p and then letting $C_f = 1/(2\pi R_2 f_p)$ gives

$$C_f = \sqrt{(1 + R_2/R_1)/f_{t1} f_{t2}}/(2\pi R_2) \tag{8.41}$$

The closed-loop bandwidth is $f_B = f_p$. It can be shown (see Problem 8.66) that increasing C_f by the factor $(1 + R_2/R_1)^{1/4}$ will make the crossover frequency f_x

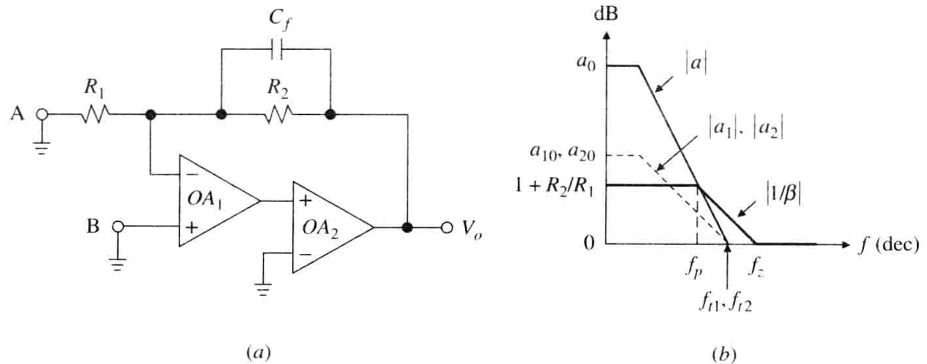


FIGURE 8.57
Composite amplifier with feedback-lead compensation.

coincide with the geometric mean $\sqrt{f_p f_z}$ and thus maximize ϕ_m; however, this will also decrease the closed-loop bandwidth in proportion.

EXAMPLE 8.17. (*a*) The circuit of Fig. 8.57*a* is to be used as a noninverting amplifier with $R_1 = 1$ kΩ and $R_2 = 99$ kΩ. (*a*) Assuming op amps with GBP = 1 MHz, find C_f for $\phi_m = 45°$. Then compare ϕ_m, T_0, and f_B with the case of a single-op amp realization. (*b*) Find C_f for the maximum phase margin. What are the resulting values of ϕ_m and f_p? (*c*) What happens if C_f is increased above the value found in (*b*)?

Solution.

(*a*) Insert the input source at node B. Letting $f_{t1} = f_{t2} = 1$ MHz in Eq. (8.41) gives $C_f = 16.1$ pF for $\phi_m = 45°$. Moreover, $T_0 = a_0^2/100 = 4 \times 10^8$, and $f_B = f_p = 100$ kHz. Had a single op amp been used, then $\phi_m = 90°$, $T_0 = a_0/100 = 2 \times 10^3$, and $f_B = 10^6/100 = 10$ kHz.

(*b*) $C_f = (100)^{1/4} \times 16.1 = 50.8$ pF, $f_p = 31.62$ kHz, $\phi_m = 180° + \measuredangle a - \measuredangle(1/\beta) \cong 180° - 180° - [\tan^{-1}(f_x/f_z) - \tan^{-1}(f_x/f_p)] = -(\tan^{-1} 0.1 - \tan^{-1} 10) = 78.6°$.

(*c*) Increasing C_f above 50.8 pF will reduce ϕ_m until eventually $\phi_m \to 0°$, indicating that overcompensation is detrimental.

In Fig. 8.57 we have stabilized the composite amplifier by acting on its feedback network. An alternative[21] type of compensation is by controlling the pole of the second op amp using local feedback, in the manner depicted in Fig. 8.58. The composite response $a = a_1 A_2$ has the dc gain $a_0 = a_{10}(1 + R_4/R_3)$, and two pole frequencies at f_{b1} and at $f_{B2} = f_{t2}/(1 + R_4/R_3)$. Without the second amplifier, the closed-loop bandwidth would be $f_{B1} = f_{t1}/(1+R_2/R_1)$. With the second amplifier in place, the bandwidth is expanded to $f_B = (1+R_4/R_3) f_{B1} = f_{t1}(1+R_4/R_3)/(1 + R_2/R_1)$. It is apparent that if we align f_B and f_{B2}, then ROC = 30 dB/dec, or $\phi_m = 45°$. Thus, imposing $f_{t1}(1 + R_4/R_3)/(1 + R_2/R_1) = f_{t2}/(1 + R_4/R_3)$ yields

$$1 + R_4/R_3 = \sqrt{(f_{t2}/f_{t1})(1 + R_2/R_1)} \tag{8.42}$$

We observe that for the benefits of using OA_2 to be significant the application must call for a sufficiently high closed-loop gain.

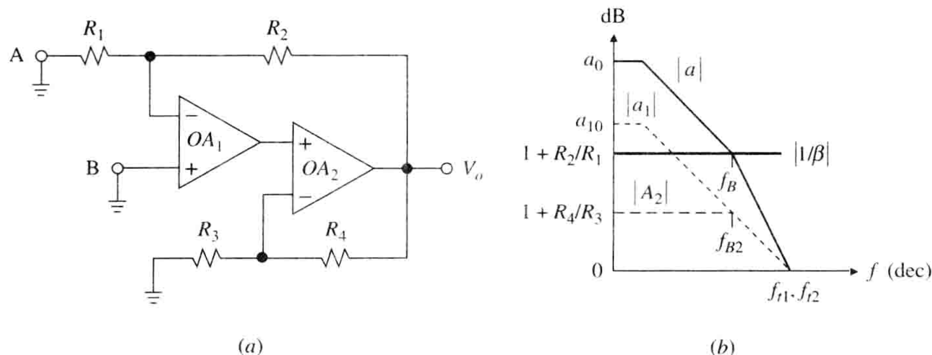


FIGURE 8.58
Composite amplifier with compensation provided by OA_2.

EXAMPLE 8.18. (*a*) Assuming op amps with GBP = 1 MHz in the circuit of Fig. 8.58*a*, specify suitable components for operation as an inverting amplifier with a dc gain of −100 V/V. Compare with a single-op-amp realization.

Solution. Insert the input source at node A and let $R_1 = 1\ \text{k}\Omega$ and $R_2 = 100\ \text{k}\Omega$. Then, $R_4/R_3 = \sqrt{101} - 1 = 9.05$. Pick $R_3 = 2\ \text{k}\Omega$ and $R_4 = 18\ \text{k}\Omega$. The dc loop gain is $T_0 = a_{10}(1 + R_4/R_3)/(1 + R_2/R_1) \cong 2 \times 10^4$, and the closed-loop bandwidth is $f_B \cong f_t/10 = 100$ kHz. If only one op amp had been used, then $\phi_m \cong 90°$, $T_0 \cong 2 \times 10^3$ and $f_B \cong 10$ kHz, indicating an order-of-magnitude improvement brought about by the second op amp.

Optimizing dc and ac Characteristics

There are applications in which it is desirable to combine the dc characteristics of a low-offset, low-noise device, such as a bipolar voltage-feedback amplifier (VFA), with the dynamics of a high-speed device, such as a current-feedback amplifier (CFA). The two sets of technologically conflicting specifications can be met with a composite amplifier. In the topology of Fig. 8.59*a* we use a CFA with local feedback to shift the $|a_1|_{\text{dB}}$ curve upward by the amount $|A_2|_{\text{dB}}$, and thus improve the dc loop gain by the same amount. As long as $f_{B2} \gg f_{t1}$, the phase shift due to the pole frequency at $f = f_{B2}$ will be insignificant at $f = f_{t1}$, indicating that we can operate the VFA with a feedback factor of unity, or at the maximum bandwidth f_{t1}. Imposing

$$1 + R_4/R_3 = 1 + R_2/R_1 \tag{8.43}$$

will maximize also the closed-loop bandwidth f_B of the composite device, which is now $f_B = f_{t1}$.

The composite topology offers important advantages other than bandwidth. Since the CFA is operated within the feedback loop of the VFA, its generally poorer input dc and noise characteristics become insignificant when referred to the input of the composite device, where they are divided by a_1. Moreover, with most of the signal swing being provided by the CFA, the slew-rate requirements of the VFA are significantly relaxed, thus ensuring high full-power bandwidth (FPB) capabilities for the composite device. Finally, since the VFA is spared from having to drive the

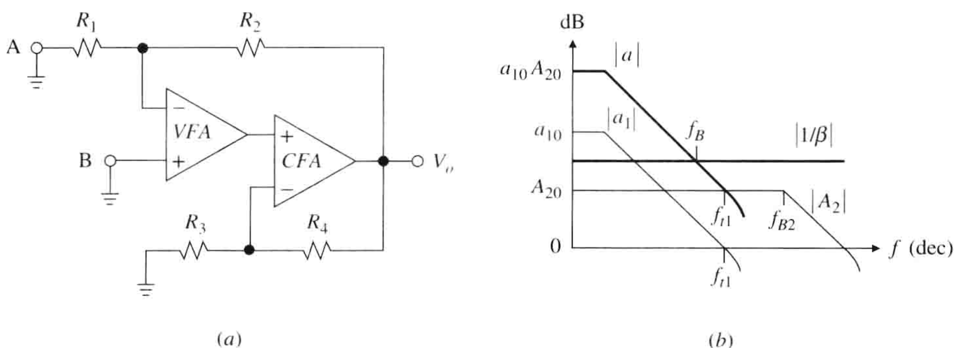


FIGURE 8.59
VFA-CFA composite amplifier.

output load, self-heating effects such as thermal feedback become insignificant, so the composite device retains optimum input-drift characteristics.

There are practical limitations to the amount of closed-loop gain achievable with a CFA. Even so, it pays to use a CFA as part of a composite amplifier. For instance, suppose we need an overall dc gain $A_0 = 10^3$ V/V, but using a CFA having only $A_{20} = 50$ V/V. Clearly, the VFA will now have to operate with a gain of $A_0/50 = 20$ V/V and a bandwidth $f_{t1}/20$. This is still 50 times better than if the VFA were to operate alone, not to mention the slew-rate and thermal-drift advantages.

In the arrangement of Fig. 8.59*a* the composite bandwidth is set by the VFA, so the amplification provided by the CFA above this band is in effect wasted. The alternative topology of Fig. 8.60 exploits the dynamics of OA_2 to their fullest extent by allowing it to participate directly in the feedback mode, but only at high frequencies. The circuit works as follows.

At dc, where the capacitances act as opens, the circuit reduces to that of Fig. 8.57*a*, so $a_0 = a_{10}a_{20}$. Clearly, the dc characteristics are set by OA_1, which provides OA_2 with whatever drive is needed to force $V_n \to V_{OS1}$. Moreover, any gross bias current at the inverting input of OA_2 is prevented from disturbing node V_n because of the dc blocking action by C_2.

As we increase the operating frequency, we witness a gradual decrease in OA_1's gain $A_1 = -1/(jf/f_1)$, $f_1 = 1/(2\pi R_3 C_1)$, while the crossover network $C_2 R_4$ gradually changes the mode of operation of OA_2 from open-loop to closed-loop. Above the crossover network frequency $f_2 = 1/(2\pi R_4 C_2)$, we can write $V_o \cong a_2(A_1 V_n - V_n)$, or

$$V_o \cong -\frac{a_{20}}{1 + jf/f_{b2}} \frac{1 + jf/f_1}{jf/f_1} V_n$$

It is apparent that if we impose $f_1 = f_{b2}$, or $R_3 C_1 = 1/(2\pi f_{b2})$, then we obtain a *pole-zero cancellation* and $V_o = -aV_n$, $a = a_{20}/(jf/f_1) = a_{20} f_{b2}/jf \cong a_2$, indicating that the high-frequency dynamics are fully controlled by OA_2.

In a practical realization the pole-zero cancellation is difficult to maintain because f_{b2} is an ill-defined parameter. Consequently, in response to an input step, the composite device will not completely stabilize until the integrator loop has settled to its final value. The resulting settling tail[17] may be of concern in certain applications.

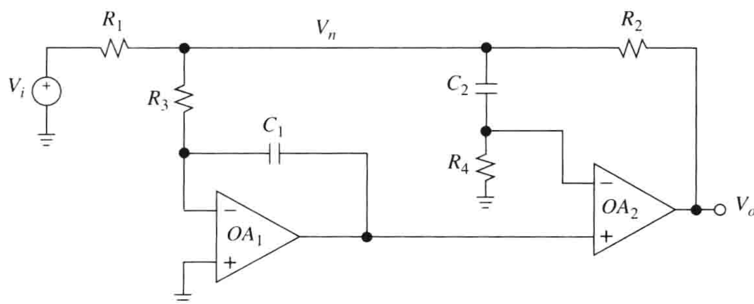


FIGURE 8.60
Composite amplifier enjoying the dc characteristics of OA_1 and the ac characteristics of OA_2.

Improving Phase Accuracy

As we know, a single-pole amplifier exhibits an error function of the type $1/(1 + 1/T) = 1/(1 + jf/f_B)$, whose phase error is $\epsilon_\phi = -\tan^{-1}(f/f_B)$, or $\epsilon_\phi \cong -f/f_B$ for $f \ll f_B$. This error is intolerable in applications requiring high phase accuracy. In the composite arrangement[22] of Fig. 8.61, OA_2 provides active feedback around OA_1 to maintain a low phase error over a much wider bandwidth than in the uncompensated case. This is similar to the active compensation of integrators of Section 6.5.

To analyze the circuit, let $\beta = R_1/(R_1 + R_2)$ and $\alpha = R_3/(R_3 + R_4)$. We note that OA_2 is a noninverting amplifier with gain $A_2 = (1/\beta)/[1 + jf/(\beta f_{t2})]$. Consequently, the feedback factor around OA_1 is $\beta_1 = \beta \times A_2 \times \alpha = \alpha/[1 + jf/(\beta f_{t2})]$.

The closed-loop gain of the composite device is $A = A_1 = a_1/(1+a_1\beta_1)$, where we are using the fact that OA_1 too is operating in the noninverting mode. Substituting $a_1 \cong f_{t1}/(jf)$ and $\beta_1 = \alpha/[1 + jf/(\beta f_{t2})]$, and letting $f_{t1} = f_{t2} = f_t$, we obtain, for $\alpha = \beta$,

$$A(jf) = A_0 \frac{1 + jf/f_B}{1 + jf/f_B - (f/f_B)^2} \tag{8.44}$$

where $A_0 = 1 + R_2/R_1$ and $f_B = f_t/A_0$. As discussed in Section 6.5, this error function offers the advantage of a very small phase error, namely, $\epsilon_\phi = -\tan^{-1}(f/f_B)^3$, or $\epsilon_\phi \cong -(f/f_B)^3$ for $f \ll f_B$.

Figure 8.62 (top) shows the results of the PSpice simulation of a composite amplifier with $A_0 = 10$ V/V using a matched pair of 10-MHz op amps, so that $f_B = 1$ MHz. For instance, at 1/10 of f_B, or 100 kHz, the composite circuit gives $\epsilon_\phi = -0.057°$, which is far better than $\epsilon_\phi = -5.7°$ for a single-op amp realization.

The stability situation, shown in Fig. 8.62 (bottom), reveals a rise in the $|1/\beta_1|$ curve because of the feedback pole introduced by OA_2 at $f = \beta f_{t2}$. This frequency is high enough not to compromise the stability of OA_1, yet low enough to cause a certain amount of gain peaking: this is the price we are paying for the dramatic improvement in the phase-error characteristic!

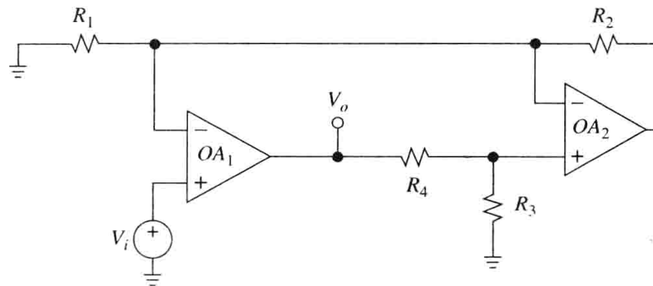


FIGURE 8.61
Composite amplifier with high phase accuracy.

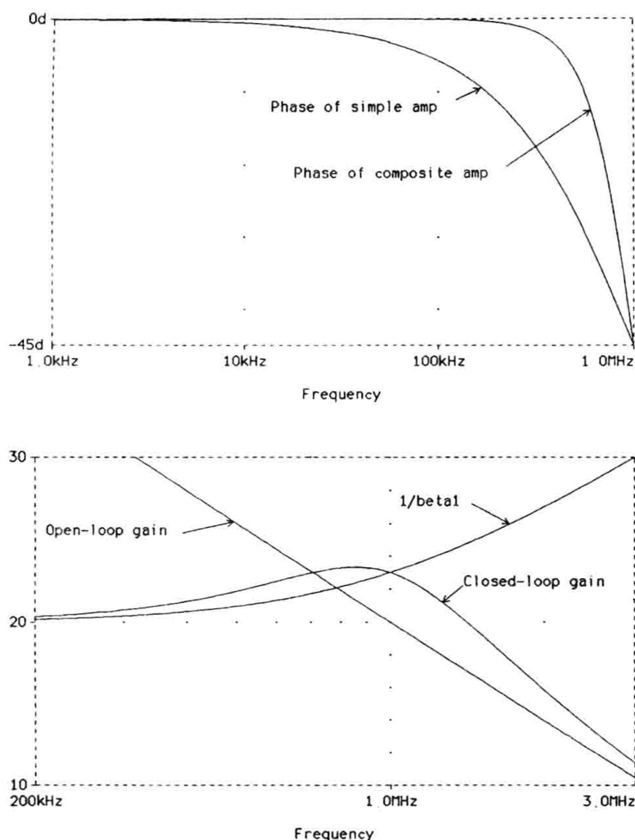


FIGURE 8.62
Frequency plots of the circuit of Fig. 8.61.

PROBLEMS

8.1 The stability problem

8.1 Suppose the loop gain of Fig. 8.1 has $T_0 = 10^4$ and three pole frequencies of 10^2 Hz, 10^6 Hz, and 10^7 Hz. (*a*) Find f_{-180°, and calculate GM. (*b*) Find f_x, and calculate ϕ_m. (*c*) Find the value to which we must lower T_0 if we want $\phi_m = 60^\circ$.

8.2 The μA702, the first monolithic op amp, had $a_0 = 3600$ V/V and three pole frequencies of 1 MHz, 4 MHz, and 40 MHz. (*a*) Find the range of values of β for which $\phi_m \geq 45^\circ$. (*b*) Find the range of values of β for which the circuit will oscillate. What is the approximate frequency of oscillation?

8.3 (*a*) What is the percentage of gain peaking at f_x of a system with $\phi_m = 30^\circ$? (*b*) What is ϕ_m for $|D(jf_x)| = 2$? For $|D(jf_x)| = 10$? For $|D(jf_x)| = 3$ dB? For $D(jf_x)| = -3$ dB? (*c*) Find β so that the circuit of Fig. 8.21 has $\phi_m = 75^\circ$. What is the value of $|D(jf_x)$?

8.4 A voltage amplifier with open-loop gain $a(jf) = a_0/[(1 + jf/f_1) \times (1 + jf/f_2)]$ is operated with $\beta = 0.1$ V/V. (*a*) If the closed-loop dc gain is $A_0 = 9$ V/V, find a_0; hence, develop a standard-form expression for $A(jf)$ in terms of f_1 and f_2. (*b*) If it is found that the phase and magnitude of $A(jf)$ at $f = 10$ kHz are $-90°$ and 90/11 V/V, what are the values of f_1 and f_2? (*c*) Find the crossover frequency f_x and, hence, the phase margin ϕ_m. (*d*) Find ϕ_m if β is raised to 1 V/V.

8.5 An amplifier with open-loop gain $a(s) = 100/[(1 + s/10^3) \times (1 + s/10^5)]$ is placed in a negative-feedback loop. (*a*) Derive an expression for the closed-loop gain $A(s)$ as a function of the feedback factor β, and find the value of β that causes the poles of $A(s)$ to be *coincident*. What is their common value? (*b*) Find the crossover frequency f_x and, hence, the phase margin ϕ_m.

8.6 An amplifier with dc gain $a_0 = 10^5$ and three pole frequencies $f_1 = 1$ kHz, $f_2 = 1$ MHz, and $f_3 = 10$ MHz is to be operated with frequency-independent feedback. (*a*) Find β for $\phi_m = 60°$. What is the corresponding value of GM? (*b*) Find β for GM = 20 dB. What is the corresponding value of ϕ_m?

8.7 An op amp with $a_0 = 10^3$ V/V and two pole frequencies at $f_1 = 100$ kHz and $f_2 = 2$ MHz is connected as a unity-gain voltage follower. Find ϕ_m, ζ, Q, GP, OS, and $A(jf)$. Would you have much use for this circuit?

8.8 An amplifier has three identical pole frequencies so that $a(jf) = a_0/(1 + jf/f_1)^3$, and is placed in a negative-feedback loop with a frequency-independent feedback factor β. Find an expression for $f_{-180°}$ as well as the corresponding value of T.

8.9 (*a*) Verify that a circuit with a dc loop gain $T_0 = 10^2$ and three pole frequencies at $f_1 = 100$ kHz, $f_2 = 1$ MHz, and $f_3 = 2$ MHz is unstable. (*b*) One way of stabilizing it is by reducing T_0. Find the value to which T_0 must be reduced for $\phi_m = 45°$. (*c*) Another way of stabilizing it is by rearranging one or more of its poles. Find the value to which f_1 must be reduced for $\phi_m = 45°$. (*d*) Repeat parts (*b*) and (*c*), but for $\phi_m = 60°$.

8.10 An amplifier with $a(jf) = 10^5(1 + jf/10^4)/[(1 + jf/10) \times (1 + jf/10^3)]$ V/V is placed in a negative-feedback loop with frequency-independent β. (*a*) Find the range of values of β for which $\phi_m \geq 45°$. (*b*) Repeat, but for $\phi_m \geq 60°$. (*c*) Find the value of β that minimizes ϕ_m. What is $\phi_{m(\min)}$?

8.11 Two negative-feedback systems are compared at some frequency f_1. If it is found that the first has $T(jf_1) = 10\,\underline{/-180°}$ and the second has $T(jf_1) = 10\,\underline{/-90°}$, which system enjoys the smaller magnitude error? The smaller phase error?

8.12 The response of a negative-feedback circuit with $\beta = 0.1$ V/V is observed with the oscilloscope. For a 1-V input step, the output exhibits an overshoot of 12.6% and a final value of 9 V. Moreover, with an ac input, the phase difference between output and input reaches 90° for $f = 10$ kHz. Assuming a second-pole error amplifier, find its open-loop response.

8.13 As mentioned, the rate-of-closure considerations hold only for minimum-phase systems. Verify by comparing the Bode plots of the minimum-phase function $H(s) = (1 + s/2\pi 10^3)/[(1 + s/2\pi 10)(1 + s/2\pi 10^2)]$ with those of the function $H(s) = (1 - s/2\pi 10^3)/[(1 + s/2\pi 10)(1 + s/2\pi 10^2)]$, which is similar to the former, except that its zero is located in the right half of the complex plane.

8.14 (*a*) An op amp with $r_i = \infty$, $r_o = 0$, and $a(jf) = 10^3/(1 + jf/10^3)$ is configured for operation as an inverting *integrator* with $f_0 = 10$ kHz. Use return-ratio analysis to calculate its phase margin. (*b*) Repeat, if $a(jf)$ has an additional pole frequency at 1 MHz. (*c*) Comment on your results.

8.15 Repeat problem 8.14 if the resistance and capacitance are interchanged with each other so as to turn the circuit into an inverting *differentiator*.

8.16 (*a*) Assuming ideal op amp, find the gain margin GM of the equal-component *KRC* filter of Example 3.8. (*b*) What happens if the op amp, instead of being ideal, has a constant GBP of 1 MHz?

8.17 Find the phase and gain margins if the circuit of Fig. 8.14 is configured as a voltage follower by letting $R_1 = \infty$ and $R_2 = 0$. Use two different injection points, and verify that even though T_v and/or T_i vary with the point of injection, T remains the same.

8.18 Suppose the high-sensitivity I-V converter of Example 2.2 is implemented with a 741 op amp with the same C_n and R_L-C_L values of Fig. 8.14. Find an injection point where only a single injection suffices, and use it to measure ϕ_m and GM. Compare the phase erosion by C_n and C_L with Example 8.2, and explain why it is less critical in the current circuit.

8.19 Figure P8.19 shows the voltage-injection method applied to some op amp circuit. (*a*) Calculate T if $a = 16$ V/V, $Z_o = Z_1 = 1$ kΩ, and $Z_2 = 2$ kΩ. (*b*) Calculate T_v. (*c*) Based on the above results, what do you expect the value of T_i to be? (*d*) Verify the above value of T_i via current injection.

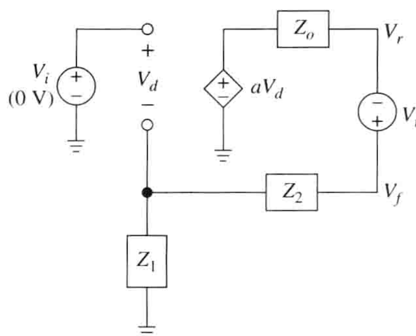


FIGURE P8.19

8.20 A certain unity-gain inverting amplifier is at the onset of peaking for having $\phi_m = \cos^{-1}(\sqrt{2} - 1)$. Assuming a_{ft} is a real number, find conditions for a_{ft} that cause peaking by making $|A(jf_x)| > 1$.

8.3 Frequency compensation of op amps

8.21 The μA702 op amp has $a_0 = 3600$ V/V, $f_1 = 1$ MHz, $f_2 = 4$ MHz, and $f_3 = 40$ MHz and is operated as a unity-gain buffer. Assuming we can change f_1 without affecting f_2 and f_3, estimate $f_{1(\text{new})}$ for a dominant-pole compensation such that (*a*) $\phi_m = 60°$, (*b*) GM = 12 dB, (*c*) GP = 2 dB, and (*d*) OS = 5%.

8.22 Consider a negative-feedback loop with dc gain $T_0 = 10^3$, a dominant pole at 1 kHz, and a pair of coincident poles at 250 kHz. (*a*) Verify that with $\beta = 1$ this loop is unstable. (*b*) To what value must f_1 be reduced for $\phi_m = 45°$? (*c*) Another way to stabilize the circuit is to reduce T_0 so that the ensuing downward shift of the magnitude plot will lower f_x to a frequency region of less phase lag. To what value must T_0 be reduced for $\phi_m = 45°$? (*d*) Repeat Parts (*b*) and (*c*), but for $\phi_m = 60°$.

8.23 An amplifier has dc gain $a_0 = 10^5$ V/V and three pole frequencies $f_1 = 100$ kHz, $f_2 = 1$ MHz, and $f_3 = 10$ MHz, arising at three nodes with equivalent resistances R_1, R_2, and R_3. (*a*) Use linearized Bode plot reasoning to verify that with $\beta = 1$ V/V this circuit is unstable. (*b*) One way of stabilizing it is to place a compensating resistance R_c in parallel with R_2 so as to raise f_2 and lower a_0. Sketch the new Bode plot if $R_c = R_2/99$, and verify that this causes a two-decade drop in a_0 as well as a two-decade increase in f_2. How do f_x and ϕ_m change? (*c*) Find the ratio R_c/R_2 that results in $\phi_m = 60°$.

8.24 Suppose in Fig. 8.21 the terminals of C_1 are inaccessible, but those of C_2 are. Repeat Example 8.5, but for C_c in parallel with C_2. Compare with the example and comment.

8.25 A certain op amp has $a_0 = 10^4$ V/V, $f_1 = 1$ kHz, and an adjustable pole frequency f_2. Find β and f_2 for a maximally flat closed-loop response with a dc gain of 60 dB. What is its -3-dB frequency?

8.26 A student wants to implement a unity-gain voltage buffer using a voltage comparator having a dc gain of 10^4 V/V, a pole frequency of 100 kHz, and a pole pair of 10 MHz. (*a*) Use ROC reasoning to verify that the circuit is unstable. (*b*) Although comparators are meant for open-loop operation, the student can still stabilize the buffer by creating a suitable *additional dominant pole* f_d via an external *R*-*C* network, as shown in Fig. P8.26. Find f_d for $\phi_m = 70°$. Hence, assuming $R = 30$ kΩ, find the required value of C. What is the expression for $T(jf)$ after compensation?

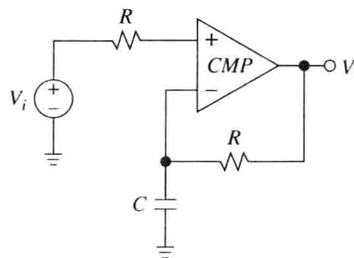


FIGURE P8.26

8.27 A student is using three identical CMOS inverters to implement a quick inverting amplifier in the manner of Fig. P8.27. Each inverter has $g_m = 1$ mA/V, $r_o = 20$ kΩ, and an equivalent capacitance $C = 1$ pF between its output node and ground. (*a*) Assuming $R_2 = 2R_1 = 200$ kΩ, use ROC reasoning to show that the circuit is unstable. (*b*) Find a capacitance C_c that, connected between the output node of I_1 and ground, will stabilize the circuit for $\phi_m = 70°$. Verify with PSpice and find the actual phase margin and the closed-loop -3-dB frequency.

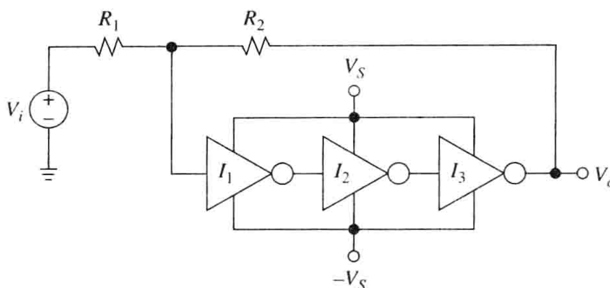


FIGURE P8.27

8.28 Obtain an expression for the gain $a_1(jf) = V_{o1}/V_i$ of the circuit of Fig. P8.28. Hence, assuming $R_c \ll R_1$ and $C_c \gg C_1$, prove Eq. (8.23). *Hint:* Given two real roots such that $\omega_a \ll \omega_b$, you can approximate $(1+s/\omega_a) \times (1+s/\omega_b) \cong 1+s/\omega_a+s^2/(\omega_a\omega_b)$.

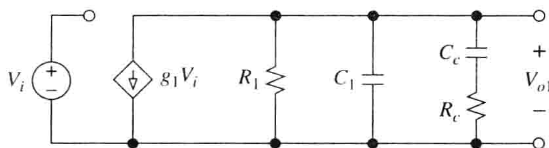


FIGURE P8.28

8.29 Specify R_c and C_c for pole-zero compensation of the circuit of Fig. 8.18 for $\phi_m \cong 60°$ with $\beta = 0.2$. Verify with PSpice.

8.30 Specify R_c and C_c for pole-zero compensation of the circuit of Problem 8.27 with $\phi_m \cong 60°$.

8.31 Derive Eqs. (8.25) and (8.26). *Hint:* Apply KCL at nodes V_1 and V_2, eliminate V_1, and use the hint of Problem 8.28.

8.32 An amplifier with a dc gain of 40,000 V/V and three pole frequencies of 100 kHz, 3 MHz, and 5 MHz is operated with $\beta = 0.5$ V/V. (*a*) Sketch and label the linearized Bode plot of $|T|$ for a visual estimation of f_x and ϕ_x. Is the circuit unstable? (*b*) Given that the first two poles are produced at the input and output nodes of an internal stage that has a dc gain of -200 V/V and input-node and output-node resistances $R_1 = 100$ kΩ and $R_2 = 10$ kΩ, find the capacitance C_c that, connected between the input and output terminals of this internal stage, stabilizes the circuit for $\phi_m = 45°$. (*c*) What are the values of this stage's g_m, C_1, and C_2? What are the values of the new pole frequencies as well as the RHP zero frequency? Use these values to write an expression for $T(jf)$.

8.33 Consider two amplifiers having, respectively, open-loop gains

$$a_1(jf) = \frac{10^4(1+jf/10^5)}{(1+jf/10^3)(1+jf/10^7)} \qquad a_2(jf) = \frac{10^4(1-jf/10^5)}{(1+jf/10^3)(1+jf/10^7)}$$

Sketch and label their linearized Bode plots (magnitude as well as phase). Find the phase margins for operation with $\beta = 1$, compare the two amplifiers, and comment on similarities as well as differences.

8.34 Use a suitable feedback network R_c-C_c across I_2 in the circuit of Problem 8.27 to provide Miller compensation for $\phi_m = 75°$ and zero removal. Verify with PSpice.

8.35 Two-stage CMOS op amps are notoriously prone to the phase erosion by the RHP zero. An alternative to using a resistor R_c to remove the offending zero is to exploit the unidirectionality of a voltage buffer, modeled in Fig. P8.35 by the $1V_o$ dependent source. The buffer retains the multiplicative action of the Miller effect while blocking forward transmission to the output, thereby removing the RHP zero. Find the gain $a(s) = V_o/V_i$, verify that the transmission zero has been moved to infinity, and obtain approximate expressions for its poles f_1 and f_2 assuming $f_1 \ll f_2$.

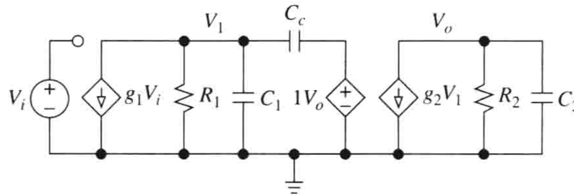


FIGURE P8.35

8.36 According to PSpice, the closed-loop gain of the feedforward-compensated amplifier of Fig. 8.32 peaks at 6.593 dB. If we were to Miller-compensate the same amplifier with a capacitance C_c across the second stage, what C_c value would yield the same closed-loop peak? Use PSpice to compare the two compensation methods, and comment.

8.4 Op amp circuits with a feedback pole

8.37 The open-loop gain of an internally compensated op amp can be approximated with a dominant pole frequency f_1 and a single high-frequency pole f_2 to account for the phase shift due to its higher-order roots. (*a*) Assuming $a_0 = 10^6$ V/V, $f_1 = 10$ Hz, and $\beta = 1$ V/V, find the actual bandwidth f_B and phase margin ϕ_m if $f_2 = 1$ MHz. (*b*) Find f_2 for $\phi_m = 60°$; what is the value of f_B? (*c*) Repeat (*b*), but for $\phi_m = 45°$.

8.38 An op amp with $a(jf) = 10^5/(1 + jf/10)$ is placed in a negative-feedback loop with $\beta(jf) = \beta_0/(1 + jf/10^5)^2$. Find the values of β_0 corresponding to (*a*) the onset of oscillatory behavior, (*b*) $\phi_m = 45°$, and (*c*) GM = 20 dB.

8.39 A Howland current pump is implemented with a constant-GBP op amp and four identical resistances. Using rate-of-closure reasoning, show that as long as the load is resistive or capacitive the circuit is stable, but can become unstable if the load is inductive. How would you compensate it?

8.40 An alternative frequency compensation method for the differentiator of Fig. 8.36*a* is by means of a suitable feedback capacitance C_f in parallel with R. Assuming $C = 10$ nF, $R = 78.7$ kΩ, and GBP = 1 MHz, specify C_f for $\phi_m = 45°$.

8.41 Show that if a constant-GBP op amp with transition frequency f_t is operated in conjunction with a feedback pole $\beta(jf) = \beta_0/(1 + jf/f_p)$, then the discrepancy function is, for $f_p \gg f_b$, $D(jf) = H_{LP}$ with $f_0 = \sqrt{\beta_0 f_t f_p}$ and $Q = \sqrt{\beta_0 f_t / f_p}$, that is, $D(jf)$ is the standard second-order low-pass function H_{LP} defined in Eq. (3.44). Clearly, the lower f_p relative to $\beta_0 f_t$, the higher the Q, which provides an alternative way for visualizing gain peaking as a consequence of inadequate phase margin.

8.42 (*a*) Use Eq. (8.1), along with the results of Problem 8.41, to prove that the transfer function of the uncompensated differentiator of Fig. 8.36*a* is $H(jf) = V_o/V_i = H_{0BP}H_{BP}(jf)$, where $H_{BP}(jf)$ is the standard second-order band-pass function defined in Eq. (3.48). (*b*) Calculate H_{0BP}, Q, the resonant frequency, and the peak value of $H(jf)$, and compare with Fig. 8.37*b*. (*c*) Repeat (*b*), but for the compensated differentiator of Fig. 8.38*b*.

8.43 Derive Eq. (8.32)

8.44 (*a*) Show that the circuit of Fig. 8.39*a* gives $A = -R_2/R_1 \times H_{LP}$, where H_{LP} is the standard second-order low-pass response defined in Eq. (3.44) with $f_0 = \sqrt{\beta_0 f_t f_z}$ and $Q = \sqrt{\beta_0 f_t/f_z}/(1 + \beta_0 f_t/f_p)$. (*b*) Find Q in the circuit of Example 8.10 before compensation. (*c*) Compensate the circuit for $\phi_m = 45°$, and find Q after compensation.

8.45 In the circuit of Example 8.10 find C_f for $\phi_m = 60°$; hence, exploit Problem 8.44 to find $A(jf)$, GP, and OS.

8.46 An alternative way of stabilizing a circuit against stray input capacitance C_n is by scaling down all resistances to raise f_z until $f_z \geq f_x$. (*a*) Scale the resistances of the circuit of Example 8.10 so that with $C_f = 0$ the circuit yields $\phi_m = 45°$. (*b*) Repeat, but for $\phi_m = 60°$. (*c*) What is the main advantage and disadvantage of this technique?

8.47 The high-sensitivity *I-V* converter of Fig. 2.2 uses $R = 1\ \text{M}\Omega$, $R_1 = 1\ \text{k}\Omega$, $R_2 = 10\ \text{k}\Omega$, and the LF351 JFET-input op amp, which has GBP = 4 MHz. (*a*) Assuming an overall input stray capacitance $C_n = 10$ pF, show that the circuit does not have enough phase margin. (*b*) Find a capacitance C_f that, when connected between the output and the inverting input, will provide neutral compensation. What is the closed-loop bandwidth of the compensated circuit?

8.48 Using an op amp with GBP = 10 MHz and $r_o = 100\ \Omega$, find the maximum C_L that can be connected to the output of the circuit of Fig. 8.45*a* and still allow for $\phi_m \geq 45°$ if (*a*) $R_1 = R_2 = 20\ \text{k}\Omega$, (*b*) $R_1 = 2\ \text{k}\Omega$, $R_2 = 18\ \text{k}\Omega$, (*c*) $R_1 = \infty$, $R_2 = 0$. (*d*) Repeat (*c*), but for $\phi_m \geq 60°$.

8.49 Using the op amp data of Problem 8.48, design an amplifier with $A_0 = +10$ V/V, under the constraint that the sum of all resistances be 200 kΩ, and that it be capable of driving a 10-nF load. Then use PSpice to verify its frequency and transient responses.

8.50 Assuming constant-GBP op amps, use linearized Bode plots to investigate the stability of (*a*) the wideband band-pass filter of Fig. 3.11, (*b*) the multiple-feedback low-pass filter of Fig. 3.32, and (*c*) the $-KRC$ low-pass filter of Problem 3.32.

8.51 (*a*) Assuming the op amp has a constant GBP of 1 MHz, discuss the stability of the multiple-feedback band-pass filter of Fig. 3.31, and verify with PSpice. (*b*) Repeat, but for the $-KRC$ band-pass filter of Problem 3.33 for the case $R_1 = R_2 = 1.607\ \text{k}\Omega$, $kR_2 = 1.445\ \text{M}\Omega$, and $C_1 = C_2 = 3.3$ nF.

8.5 Input-Lag and Feedback-Lead Compensation

8.52 Assuming the op amp of Fig. 8.47*a* has $r_d = \infty$ and $r_o = 0$, drive an expression for the noise gain $1/\beta(jf)$. Provide individual expressions for $1/\beta_0$, $1/\beta_\infty$, f_p, and f_z.

8.53 An op amp with $a_0 = 10^5$ V/V, $f_1 = 10$ kHz, $f_2 = 3$ MHz, and $f_3 = 30$ MHz, is to be used as an inverting amplifier with two 20-kΩ resistances. Use input-lag compensation to stabilize it for $\phi_m = 45°$. Hence, obtain an expression for $A(jf)$.

8.54 Use input-lag compensation to stabilize the voltage comparator of Problem 8.26 for voltage-buffer operation with $\phi_m = 65°$.

8.55 In Fig. P8.55 let $R_1 = R_2 = R_4 = 100$ kΩ, $R_3 = 10$ kΩ, and let the op amp have $a_0 = 10^5$ V/V, $f_1 = 10$ kHz, $f_2 = 200$ kHz, and $f_3 = 2$ MHz. (*a*) Verify that the circuit is unstable. (*b*) Use input-lag compensation to stabilize it for $\phi_m = 45°$. (*c*) Find the closed-loop bandwidth after compensation.

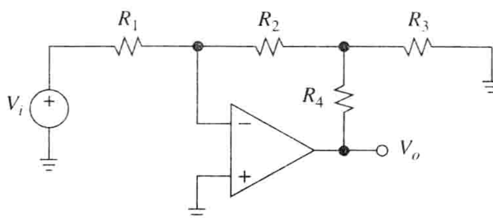


FIGURE P8.55

8.56 A certain decompensated op amp has GBP = 25 MHz for $\beta_{\max} = 0.1$ V/V, and is to be used to implement an I-V converter with a sensitivity of 0.1 V/μA. A designer is considering two alternatives: one using a 100-kΩ feedback resistor and input-lag compensation, and the other using a T-network in such a way as to make compensation unnecessary. Show the two circuit designs, and compare them in terms of closed-loop bandwidth, output error due to V_{OS}, and dc gain error.

8.57 The OPA637 op amp of Fig. P8.57 is a decompensated amplifier with SR = 135 V/μs and GBP = 80 MHz for $1/\beta \geq 5$ V/V. Since the op amp is not compensated for unity-gain stability, the integrator shown would be unstable. (*a*) Show that the circuit can be stabilized by connecting a compensation capacitance C_c as shown, and find a suitable value for C_c for $\phi_m = 45°$. (*b*) Obtain an expression for $H(jf)$ after compensation and indicate the frequency range over which the circuit behaves reasonably well as an integrator.

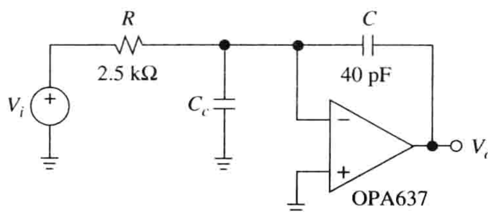


FIGURE P8.57

8.58 An op amp with GBP = 6 MHz and $r_o = 30$ Ω is to operate as a unity-gain voltage follower with an output load of 5 nF. Design an input-lag network to stabilize it. Then verify its frequency and transient responses via PSpice.

8.59 Using a decompensated op amp with GBP = 80 MHz and $\beta_{\max} = 0.2$ V/V, design a unity-gain inverting amplifier, and find $A(jf)$.

8.60 An op amp with $a_0 = 10^6$ V/V and two coincident pole frequencies $f_1 = f_2 = 10$ Hz is configured as an inverting amplifier with $R_1 = 1$ kΩ and $R_2 = 20$ kΩ. (*a*) Use feedback-lead compensation to stabilize it for $\phi_m = 45°$; then find $A(jf)$. (*b*) Find the value of C_f that will maximize ϕ_m; next find ϕ_m as well as the corresponding closed-loop bandwidth.

8.61 The wideband band-pass filter of Example 3.5 is implemented with an op amp having $a_0 = 10^5$ V/V and two pole frequencies $f_1 = 10$ Hz and $f_2 = 2$ MHz. Sketch the Bode plots of $|a|$ and $|1/\beta|$ in the vicinity of f_x and find ϕ_m.

8.6 Stability in CFA circuits

8.62 The CFA integrator of Fig. P8.62 uses a series resistance R_2 between the summing junction and the inverting-input pin to ensure $1/\beta \geq (1/\beta)_{\min}$ over frequency and thus avoid instability problems. (*a*) Investigate the stability of the circuit using Bode plots. (*b*) Assuming the CFA parameters of Problem 6.66, specify suitable components for $f_0 = 1$ MHz. (*c*) List possible disadvantages of this circuit.

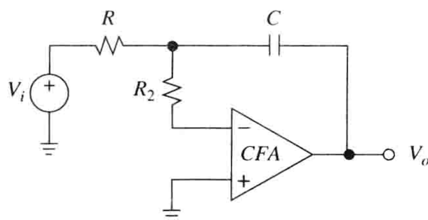


FIGURE P8.62

8.63 The CFA of Problem 6.66 is to be used to design a Butterworth band-pass filter with $f_0 = 10$ MHz and $H_{0BP} = 0$ db, and two alternatives are being considered, namely, the multiple-feedback and the *KRC* designs. Which configuration are you choosing, and why? Show the final circuit.

8.64 (*a*) Show that without C_f the CFA *I-V* converter of Fig. 8.54 yields $V_o/I_i = RH_{\text{LP}}$, where H_{LP} is the standard second-order low-pass response defined in Eq. (3.44) with $f_0 = [z_0 f_b/(2\pi r_n R_2 C_n)]^{1/2}$ and $Q = z_0 f_b/[(r_n + R_2 f_0)]$. (*b*) Predict the GP and OS for the circuit of Example 8.16 before compensation.

8.65 A certain CFA has $r_n = 50$ Ω and an open-loop dc gain of 1 V/μA, and its frequency response can be approximated with two pole frequencies, one at 100 kHz and the other at 100 MHz. The CFA is to be used as a unity-gain voltage follower. (*a*) Find the feedback resistance needed for a phase margin of 45°; what is the closed-loop bandwidth? (*b*) Repeat, but for a 60° margin.

8.7 Composite amplifiers

8.66 (*a*) With reference to the circuit of Fig. 8.57*a*, show that ϕ_m is maximized for $C_f = (1 + R_2/R_1)^{3/4}/[2\pi R_2(f_{t1} f_{t2})^{1/2}]$. (*b*) Show that for $\phi_{m(\max)} \geq 45°$ we must have

$1 + R_2/R_1 \geq \tan^2 67.5^\circ = 5.8$. (*c*) Assuming 1-MHz op amps, specify suitable component values for operation as an inverting amplifier with $A_0 = -10$ V/V and maximum phase margin. Hence, find the actual values of ϕ_m and $A(jf)$.

8.67 (*a*) Compare the circuit of Example 8.15 with a circuit implemented by cascading two amplifiers with individual dc gains $A_{10} = A_{20} = \sqrt{|A_0|}$ V/V. (*b*) Repeat, but for the circuit of Example 8.16.

8.68 An alternative to Eq. (8.42) is $1 + R_4/R_3 = \sqrt{(1 + R_2/R_1)/2}$, where we have assumed $f_{t1} = f_{t2}$. (*a*) Verify that this alternative yields $\phi_m \cong 65^\circ$. (*b*) Apply it to the design of a composite amplifier with dc gain $A_0 = -50$ V/V. (*c*) Assuming $f_{t1} = f_{t2} = 4.5$ MHz, find $A(jf)$.

8.69 In the composite amplifier of Fig. 8.60 assume OA_1 has $a_{10} = 100$ V/mV, $f_{t1} = 1$ MHz, $V_{OS1} \cong 0$, and $I_{B1} \cong 0$, and OA_2 has $a_{20} = 25$ V/mV, $f_{t2} = 500$ MHz, $V_{OS2} = 5$ mV, and $I_{B2} = 20\ \mu$A. Specify suitable components for $A_0 = -10$ V/V, under the constraint $f_2 = 0.1 f_1$. What is the output dc error E_O and the closed-loop bandwidth f_B?

8.70 For the circuit of Problem 8.69 find the total rms output noise E_{no} if $e_{n1} = 2$ nV/$\sqrt{\text{Hz}}$, $i_{n1} = 0.5$ pA/$\sqrt{\text{Hz}}$, $e_{n2} = 5$ nV/$\sqrt{\text{Hz}}$, and $i_{n2} = 5$ pA/$\sqrt{\text{Hz}}$. Ignore $1/f$ noise. Can you reduce E_{no}?

8.71 (*a*) Find ϕ_m, GP, and OS for the composite amplifier of Fig. 8.61. (*b*) Find its 1° phase-error bandwidth, and compare it with that of a single-op-amp realization with the same value of A_0, as well as with that of the cascade realization of two amplifiers with individual dc gains $\sqrt{A_0}$.

8.72 The active-compensation scheme of Fig. P8.72 (see *IEEE Trans. Circuits Syst.*, Vol. CAS-26, February 1979, pp. 112–117) works for both the inverting and the noninverting mode of operation of OA_1. Show that $V_o = [(1/\beta)V_2 + (1 - 1/\beta)V_1]/(1 + 1/T)$, $1/(1 + 1/T) = (1 + jf/\beta_2 f_{t2})/(1 + jf/\beta f_{t1} - f^2/\beta f_{t1}\beta_2 f_{t2})$, $\beta = R_1/(R_1 + R_2)$, $\beta_2 = R_3/(R_3 + R_4)$.

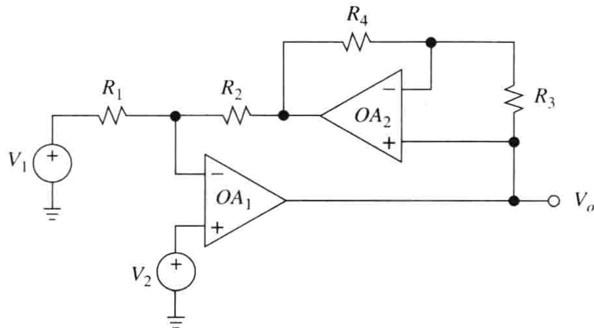


FIGURE P8.72

8.73 Apply the scheme of Problem 8.72 to the design of a high-phase-accuracy (*a*) voltage follower, (*b*) *I-V* converter with a sensitivity of 10 V/mA, and (*c*) difference amplifier with a dc gain of 100 V/V. Assume matched op amps with $f_t = 10$ MHz.

REFERENCES

1. J. K. Roberge, *Operational Amplifiers: Theory and Practice,* John Wiley & Sons, New York, 1975.
2. S. Rosenstark, *Feedback Amplifiers Principles,* Macmillan, New York, 1986.
3. R. C. Dorf and R. H. Bishop, *Modern Control Systems*, 12th ed., Prentice Hall, Upper Saddle River, NJ, 2011.
4. R. D. Middlebrook, "Measurement of Loop Gain in Feedback Systems," *Int. J. Electronics*, Vol. 38, No. 4, April 1975, pp. 485–512.
5. P. J. Hurst, "A Comparison of Two Approaches to Feedback Circuit Analysis," *IEEE Trans. on Education*, Vol. 35, No. 3, August 1992, pp. 253–261.
6. R. D. Middlebrook, "The General Feedback Theorem: A Final Solution for Feedback Systems," *IEEE Microwave Magazine*, April 2006, pp. 50–63.
7. M. Tian, V. Visvanathan, J. Hantgan, and K. Kundert, "Striving for Small-Signal Stability", *IEEE Circuits and Devices Magazine*, Vol. 17, No. 1, January 2001, pp. 31–41.
8. P. R. Gray, P. J. Hurst, S. H. Lewis, and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 5th ed., John Wiley & Sons, New York, 2009.
9. S. Franco, *Analog Circuit Design—Discrete and Integrated,* McGraw-Hill, New York, 2014.
10. J. G. Graeme, "Phase Compensation Counteracts Op Amp Input Capacitance," *EDN,* Jan. 6, 1994, pp. 97–104.
11. S. Franco, "Simple Techniques Provide Compensation for Capacitive Loads," *EDN,* June 8, 1989, pp. 147–149.
12. Dr. Alf Lundin, private correspondence, September 2003.
13. J. Williams, "High-Speed Amplifier Techniques," Application Note AN-47, *Linear Applications Handbook Volume II,* Linear Technology, Milpitas, CA, 1993.
14. A. P. Brokaw, "An IC Amplifiers User's Guide to Decoupling, Grounding, and Making Things Go Right for a Change," Application Note AN-202, *Applications Reference Manual,* Analog Devices, Norwood, MA, 1993.
15. A. P. Brokaw, "Analog Signal Handling for High Speed and Accuracy," Application Note AN-342, *Applications Reference Manual,* Analog Devices, Norwood, MA, 1993.
16. J.-H. Broeders, M. Meywes, and B. Baker, "Noise and Interference," *1996 Design Seminar,* Burr-Brown, Tucson, AZ, 1996.
17. S. Franco, "Demystfying Pole-Zero Doublets," EDN, Aug. 27, 2013 http://www.edn.com/electronics-blogs/analog-bytes/4420171/Demystifying-pole-zero-doublets.
18. Based on the author's article "Current-Feedback Amplifiers Benefit High-Speed Designs," *EDN,* Jan. 5, 1989, pp. 161–172. © Cahners Publishing Company, 1997, a Division of Reed Elsevier Inc.
19. J. Williams, "Composite Amplifiers," Application Note AN-21, *Linear Applications Handbook Volume I,* Linear Technology, Milpitas, CA, 1990.
20. J. Graeme, "Phase Compensation Perks Up Composite Amplifiers," *Electronic Design,* Aug. 19, 1993, pp. 64–78.
21. J. Graeme, "Composite Amplifier Hikes Precision and Speed," *Electronic Design Analog Applications Issue,* June 24, 1993, pp. 30–38.
22. J. Wong, "Active Feedback Improves Amplifier Phase Accuracy," *EDN,* Sept. 17, 1987.

9

NONLINEAR CIRCUITS

到目前为止，我们所讨论的所有电路都是在线性条件下工作的。线性可以通过（a）使用负反馈来强制使得运放工作在其线性区以及（b）使用线性元件来实现反馈网络来实现。

使用一个正反馈的高增益放大器，或者根本不使用反馈，会使得器件基本上工作在饱和区。这种双稳态特性具有高度的非线性并且是电压比较器和施密特触发器的基础。

非线性特性也能够通过在反馈网络中使用非线性元件（如二极管或模拟开关）来实现。常见的例子包括精密整流器、峰值检测器以及采样保持放大器。另外一类非线性电路利用 BJT 可以预估的指数特性来实现不同的非线性传递特性，例如对数放大和模拟乘法。这一类非线性电路将在第 13 章详细讨论。

本章重点

本章从电压比较器及其性能特性（如响应时间和逻辑电平）开始。随后介绍一些常见的应用：电平检测、开 / 关控制、窗口比较器、条形图计以及脉宽调制。比较器的部分总结了施密特触发器及其在消抖和滞回开关控制中的应用。

本章第二部分包含基于二极管的非线性电路及其应用，如超级二极管、半波和全波整流电路以及交流 - 直流转换器。

本章最后介绍模拟开关、峰值检测器以及采样保持放大器。

9.1 VOLTAGE COMPARATORS

The function of a voltage comparator is to compare the voltage v_P at one of its inputs against the voltage v_N at the other, and output either a low voltage V_{OL} or a high voltage V_{OH} according to

$$v_O = V_{OL} \qquad \text{for } v_P < v_N \tag{9.1a}$$

$$v_O = V_{OH} \qquad \text{for } v_P > v_N \tag{9.1b}$$

As shown in Fig. 9.1*a*, the symbolism used for comparators is the same as for op amps. We observe that while v_P and v_N are *analog* variables because they can assume a continuum of values, v_O is a *binary* variable because it can assume only one of two values, V_{OL} or V_{OH}. It is fair to view the comparator as a one-bit analog-to-digital converter.

Introducing the differential input voltage $v_D = v_P - v_N$, the above equations can also be expressed as $v_O = V_{OL}$ for $v_D < 0$ V, and $v_O = V_{OH}$ for $v_D > 0$ V. The voltage transfer curve (VTC), shown in Fig. 9.1*b*, is a nonlinear curve. At the origin, the curve is a vertical segment, indicating an infinite gain there, or $v_O/v_D = \infty$. A practical comparator can only approximate this idealized VTC, with actual gains

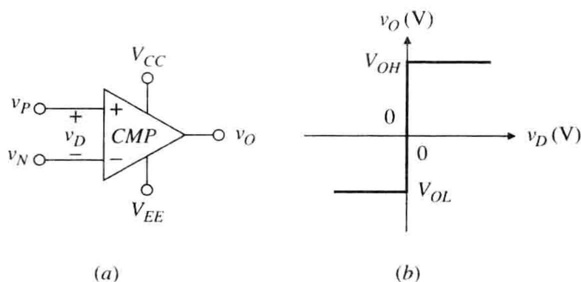


FIGURE 9.1
Voltage-comparator symbolism and ideal VTC. (All node voltages are referenced to ground.)

being typically in the range from 10^3 to 10^6 V/V. Away from the origin, the VTC consists of two horizontal lines positioned at $v_O = V_{OL}$ and $v_O = V_{OH}$. These levels need not necessarily be symmetric, though symmetry may be desirable in certain applications. All that matters is that the two levels be sufficiently far apart to make their distinction reliable. For example, digital applications require $V_{OL} \cong 0$ V and $V_{OH} \cong 5$ V.

The Response Time

In high-speed applications it is of interest to know how rapidly a comparator responds as the input state changes from $v_P < v_N$ to $v_P > v_N$, and vice versa. Comparator speed is characterized in terms of the *response time,* also called the *propagation delay* t_{PD}, defined as the time it takes for the output to accomplish 50% of its transition in response to a predetermined voltage step at the input. Figure 9.2 illustrates the setup for the measurement of t_{PD}. Though the input step magnitude is typically on the order of 100 mV, its limits are chosen to barely exceed the level required to cause the output to switch states. This excess voltage is called the *input overdrive* V_{od}, with typical overdrive values being 1 mV, 5 mV, and 10 mV. In general t_{PD} decreases with V_{od}. Depending on the particular device and the value of V_{od}, t_{PD} can range from a few microseconds to a few nanoseconds.

The Op Amp as a Voltage Comparator

When speed is not critical, an op amp can make an excellent comparator,[1] especially in view of the extremely high gains and low input offsets available from many popular op amp families. The VTC of a practical op amp was depicted in Fig. 1.46, where we expressed v_D in microvolts in order to be able to visualize the slope of the VTC in the linear region. In comparator applications v_D can be a hefty signal, so it is more

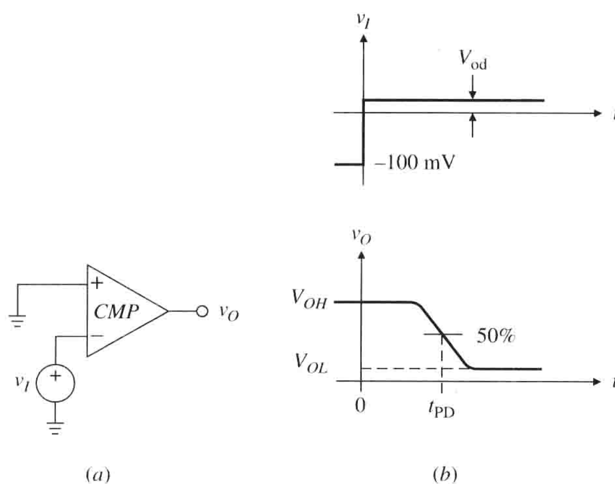


FIGURE 9.2
The response time of a comparator.

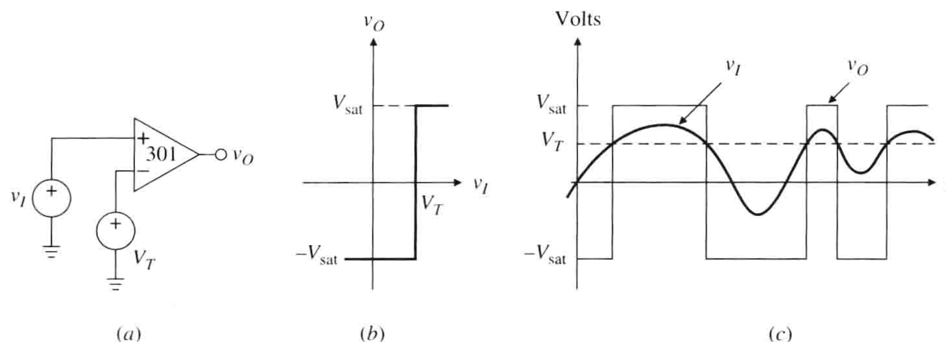


FIGURE 9.3
Threshold detector.

appropriate to express it in volts than in microvolts. If we do so, the horizontal scale undergoes so much compression that the linear-region portion of the VTC coalesces with the vertical axis, resulting in a curve of the type of Fig. 9.1*b*.

The circuit of Fig. 9.3*a* uses a 301 op amp to compare v_I against some voltage threshold, V_T. When $v_I < V_T$ the circuit gives $v_O = -V_{sat} \cong -13$ V, and when $v_I > V_T$ it gives $v_O = +V_{sat} \cong +13$ V. This is illustrated in the figure via both the VTC and the voltage waveforms. Since v_O goes high whenever v_I rises above V_T, the circuit is aptly called a *threshold detector.* If $V_T = 0$ V, the circuit is referred to as a *zero-crossing detector.*

It is important to realize that when used as a comparator, the op amp has no control over v_N due to the absence of feedback. The amplifier now operates in the open-loop mode and, because of its extremely high gain, it spends most of its time in saturation. Clearly, v_N no longer tracks v_P!

Though the output transitions in Fig. 9.3*c* have been shown as instantaneous, we know that in practice they take some time due to slew-rate limiting. Had we used a 741 op amp, the time to accomplish 50% of the output transition would have been $t_R = V_{sat}/\text{SR} = (13\text{ V})/(0.5\text{ V}/\mu\text{s}) = 26\ \mu\text{s}$, an intolerably long time in many applications. The reason for using the 301 op amp is that it comes without the internal frequency-compensation capacitance C_c, so it slews more rapidly than the 741 op amp. Frequency compensation is indispensable in negative-feedback applications but is superfluous in open-loop applications, where it only slows down the comparator unnecessarily.

Whether internally compensated or not, op amps are intended for negative-feedback operation, so their dynamics are not necessarily optimized for open-loop operation. Moreover, their output saturation levels are generally awkward to interface to digital circuitry. These and other needs peculiar of the voltage-comparison operation have provided the motivation for developing a category of high-gain amplifiers specifically optimized for this operation and thus called voltage comparators.

General-Purpose IC Comparators

Figure 9.4 depicts one of the earliest and most popular voltage comparators, the LM311. The input stage consists of the *pnp* emitter followers Q_1 and Q_2 driving the

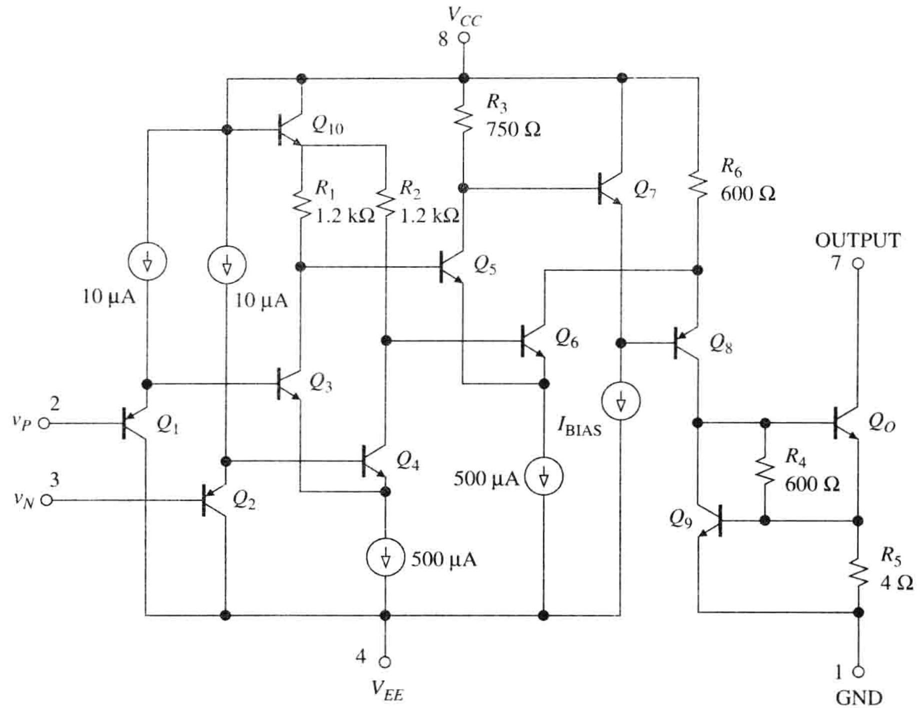


FIGURE 9.4
Simplified circuit diagram of the LM311 voltage comparator. (Courtesy of Texas Instruments.)

differential pair Q_3-Q_4. The output of this pair is further amplified by the Q_5-Q_6 pair and then by the Q_7-Q_8 pair, from which it emerges as a single-ended current drive for the base of the output transistor Q_O. Circuit operation is such that for $v_P < v_N$, Q_8 sources substantial current to the base of Q_O, keeping it in heavy conduction; for $v_P > v_N$, the base drive is removed and Q_O is thus in cutoff. Summarizing,

$$Q_O = \text{Off} \qquad \text{for } v_P > v_N \tag{9.2a}$$

$$Q_O = \text{On} \qquad \text{for } v_P < v_N \tag{9.2b}$$

The function of Q_9 and R_5 is to provide overload protection for Q_O, in the manner discussed in Section 5.8 for op amps.

When on, Q_O can draw up to 50 mA of current. When off, it draws a negligible leakage current of 0.2 nA typical. Both the collector and the emitter terminals (ignoring R_5) are externally accessible to allow for custom biasing of Q_O. The most common biasing scheme involves a mere pullup resistance R_C, as shown in Fig. 9.5*a*. For $v_P < v_N$, Q_O saturates and is thus modeled with a source $V_{CE(\text{sat})}$ as in Fig 9.5*b*. So, $v_O = V_{EE(\text{logic})} + V_{CE(\text{sat})}$. Typically $V_{CE(\text{sat})} \cong 0.1$ V, so we can approximate

$$v_O = V_{OL} \cong V_{EE(\text{logic})} \qquad \text{for } v_P < v_N \tag{9.3a}$$

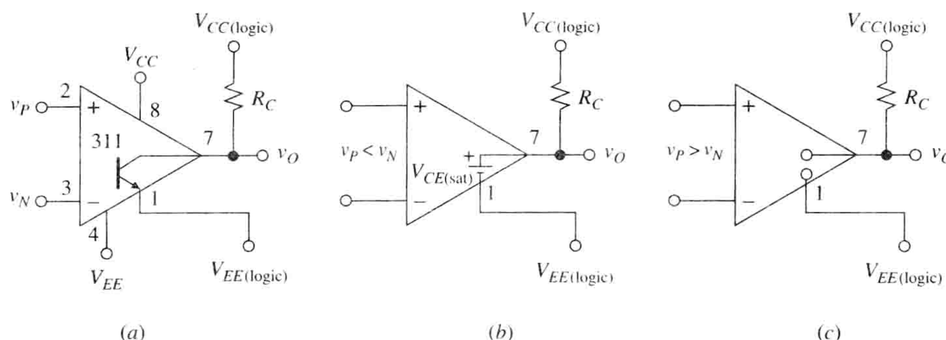


FIGURE 9.5
(*a*) Biasing the LM311 output stage with a pullup resistance R_C. Equivalent circuits for the (*b*) "output low" and (*c*) "output high" states.

For $v_P > v_N$, Q_O is in cutoff and is modeled with an open circuit as in Fig. 9.5*c*. By the pullup action of R_c we can write

$$v_O = V_{OH} \cong V_{CC(\text{logic})} \qquad \text{for } v_P > v_N \tag{9.3b}$$

The above expressions indicate that the output logic levels are under the control of the user. For example, letting $V_{CC(\text{logic})} = 5$ V and $V_{EE(\text{logic})} = 0$ V provides TTL and CMOS compatibility. Letting $V_{CC(\text{logic})} = 15$ V and $V_{EE(\text{logic})} = -15$ V yields ±15-V output levels, but without the notorious uncertainties of op amp saturation voltages. The 311 can also operate from a single 5-V logic supply if we let $V_{CC(\text{logic})} = V_{CC} = 5$ V and $V_{EE(\text{logic})} = V_{EE} = 0$ V. In fact, in the single-supply mode the device is rated to function all the way up to $V_{CC} = 36$ V.

Figure 9.6*a* shows another popular biasing scheme, which uses a pulldown resistance R_E to operate Q_O as an emitter follower. This alternative is useful when interfacing to grounded loads such as silicon-controlled rectifiers (SCRs), an example of which will be discussed in Section 11.5. The VTCs for the two biasing schemes are shown in Fig. 9.6*b*. Note the opposing polarities of the two curves.

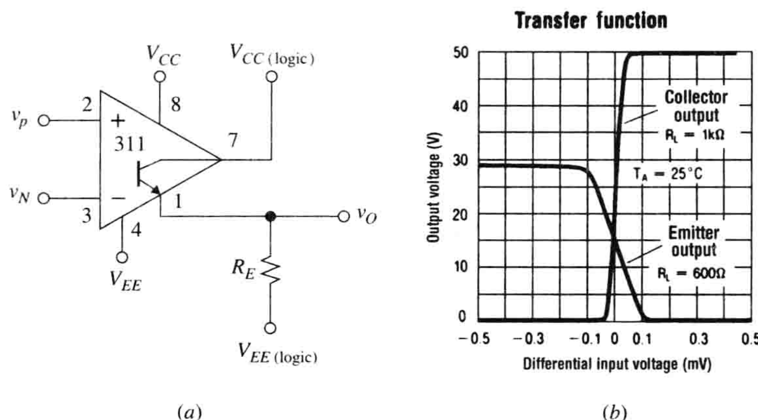


FIGURE 9.6
(*a*) Biasing the LM311 output stage with a pulldown resistance R_E. (*b*) VTC comparison for pullup and pulldown biasing. (Courtesy of Texas Instruments.)

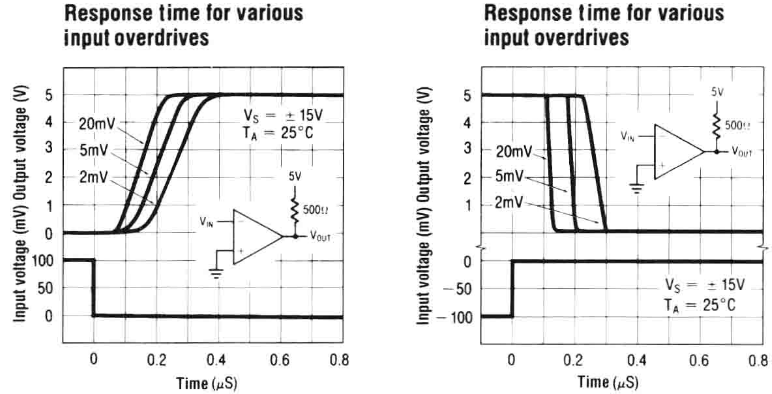


FIGURE 9.7
Typical response times of the LM311 comparator. (Courtesy of Texas Instruments.)

Figure 9.7 shows the response times of the 311 for various input overdrives. The responses corresponding to $V_{\text{od}} = 5$ mV are often used for comparing different devices. Based on the diagrams, we can characterize the 311 as basically a 200-ns comparator when used with a pullup resistor on the order of a few kilo-ohms.

Like their op amp cousins, voltage comparators suffer from dc input errors whose effect is to shift the input tripping point by an error

$$E_I = V_{OS} + R_n I_N - R_p I_P \tag{9.4}$$

where V_{OS} is the input offset voltage, I_N and I_P the currents into the inverting- and noninverting-input pins, and R_n and R_p the external dc resistances seen by the same pins. A web search for the LM311 data sheets will reveal typical room-temperature characteristics of $V_{OS} = 2$ mV, $I_B = 100$ nA (flowing out of the input pins because of the *pnp* input BJTs), and $I_{OS} = 6$ nA. The data sheets show also a scheme for externally nulling the input offset error.

Another very popular comparator, especially in low-cost single-supply applications, is the LM339 quad comparator and its derivatives. As shown in Fig. 9.8, its differential input stage is implemented with the *pnp* Darlington pairs Q_1-Q_2 and Q_3-Q_4, which result in a low-input-bias current as well as an input voltage range extending all the way down to 0 V (by contrast, the input voltage range of the LM311 extends only down to $V_{EE} + 0.5$ V, that is, within 0.5-V of the negative supply rail). The current mirror Q_5-Q_6 forms an active load for this stage and also converts to a single-ended drive for Q_7. This transistor provides additional gain as well as the base drive for the open-collector output transistor Q_O. The state of Q_O is controlled by v_P and v_N according to Eq. (9.2). Open-collector output stages are suited to wired-OR operation, just like open-collector TTL gates. When on, Q_O can sink 16 mA typical, 6 mA minimum; when off, its collector leakage is typically 0.1 nA.

The other pertinent characteristics are, typically, $V_{OS} = 2$ mV, $I_B = 25$ nA, and $I_{OS} = 5$ nA. Moreover, the operating supply range is from 2 V to 36 V, and the input voltage range is from 0 V to $V_{CC} - 1.5$ V.

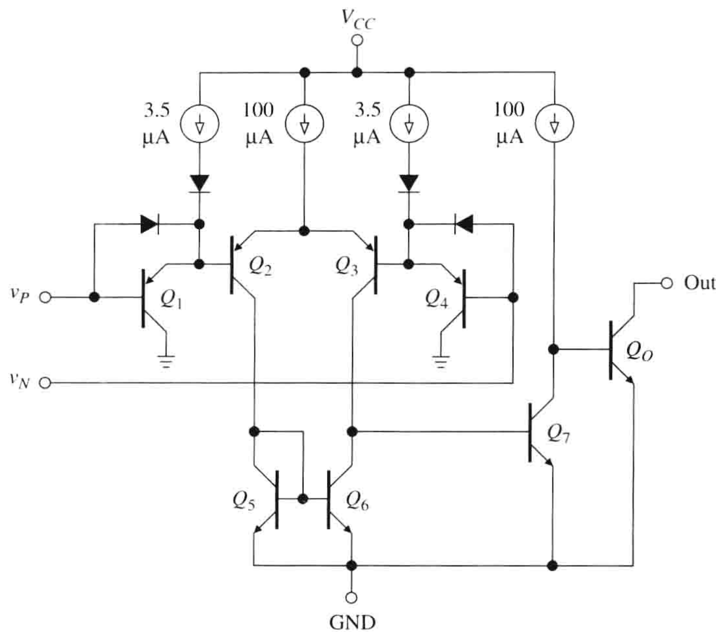


FIGURE 9.8
Simplified circuit diagram of the LM339 quad comparator. (Courtesy of Texas Instruments.)

Comparators are available in a variety of versions, such as duals and quads, low-power versions, FET-input versions, and rail-to-rail versions. The LMC7211 is a micropower CMOS comparator with rail-to-rail capabilities both at the input and at the output; the LMC7221 is similar, but with an open-drain output.

High-Speed Comparators

High-speed data converters, such as flash A-D converters, to be studied in Chapter 12, rely on the use of commensurately fast voltage comparators. To serve this and similar needs, very high-speed comparators are available with response times on the order of 10 ns or less. Such speeds are achieved through circuit techniques and fabrication processes similar to those of the faster logic families such as Schottky TTL and ECL. Moreover, to fully realize these capabilities, suitable circuit construction techniques and power-supply bypass are mandatory on the part of the user.[2]

These comparators are often equipped with output latch capabilities, which allow freezing the output state in a latch flip-flop and holding it indefinitely until the arrival of a new latch-enable command. This feature is especially useful in flash A-D converters. The symbolism and timing for these comparators are shown in Fig. 9.9. To guarantee proper output data, v_D must be valid at least t_S ns before the latch-enable command is asserted, and must remain valid for at least t_H ns thereafter, where t_S and t_H represent, respectively, the *setup* and *hold* times. Popular examples of latch comparators are the CMP-05 and LT1016. The latter has $t_S = 5$ ns, $t_H = 3$ ns, and $t_{PD} = 10$ ns.

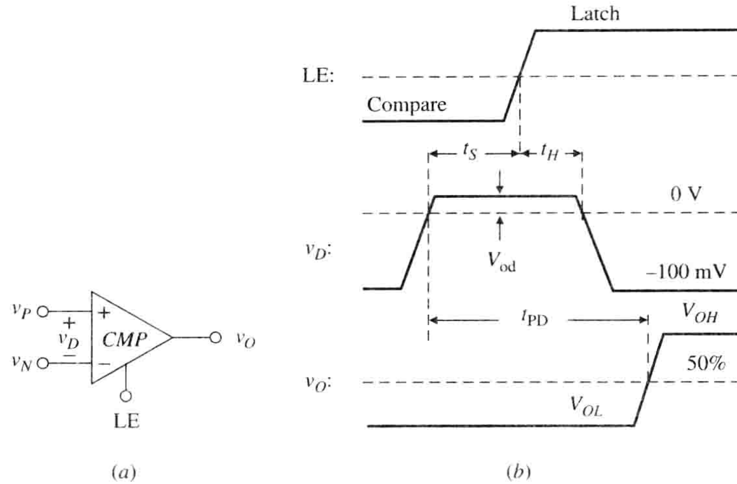


FIGURE 9.9
Comparator with latch enable, and waveforms.

Another useful feature available in some comparators is the strobe control, which disables the device by forcing its output stage into a high-impedance state. This feature is designed to facilitate bus interfacing in microprocessor applications. Finally, for increased flexibility, some comparators provide the output both in true (Q) and in negated ($\overline{Q}$) form.

SPICE Simulation of Comparators

Like their op amp cousins, voltage comparators are most efficiently simulated by means of SPICE macromodels. The PSpice circuit of Fig. 9.10*a* uses a 339 macromodel downloaded from the web to display the voltage transfer curve (VTC) of Fig. 9.10*b* (note the 28-μV offset of this particular macromodel). Likewise, we can

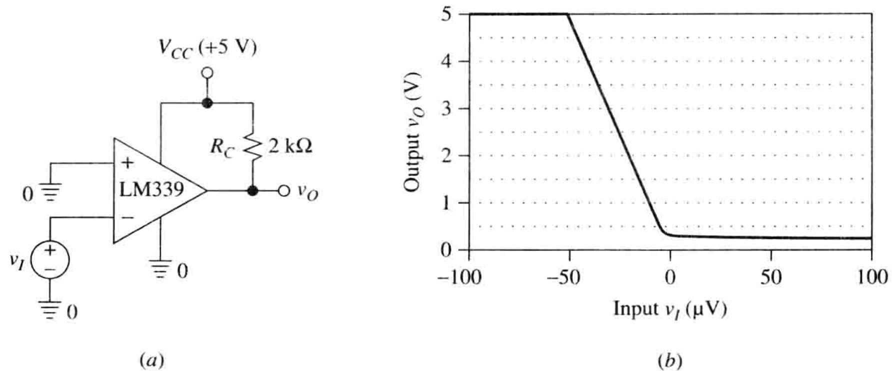


FIGURE 9.10
(*a*) PSpice circuit to display (*b*) the VTC of the 339 voltage comparator.

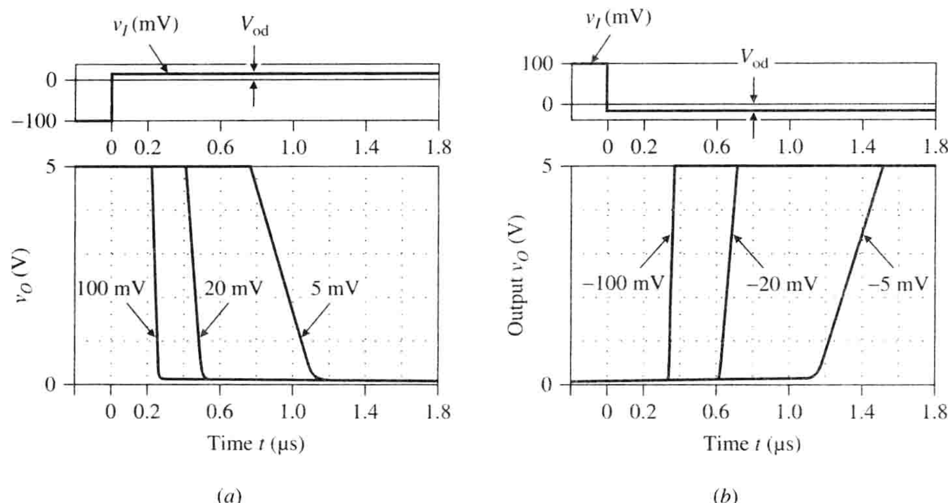


FIGURE 9.11
Response times of the 339 voltage comparator of Fig. 9.10*a* for different input overdrives.

use the circuit of 9.10*a* also to display the response times for different overdrives, as illustrated in Fig. 9.11. As in the case of the 311 comparator, the higher the overdrive, the shorter the propagation delays.

9.2 COMPARATOR APPLICATIONS

Comparators are used in various phases of signal generation and transmission, as well as in automatic control and measurement. They appear both alone or as part of systems, such as A-D converters, switching regulators, function generators, *V-F* converters, power-supply supervisors, and a variety of others.

Level Detectors

The function of a level detector, also called a *threshold detector,* is to monitor a physical variable that can be expressed in terms of a voltage, and signal whenever the variable rises above (or drops below) a prescribed value called the *set value.* The detector output is then used to undertake a specific action as demanded by the application. Typical examples are the activation of a warning indicator, such as a light-emitting diode (LED) or a buzzer, the turning on of a motor or heater, or the generation of an interrupt to a microprocessor.

As shown in Fig. 9.12, the basic components of a level detector are: (*a*) a voltage reference V_{REF} to establish a stable threshold, (*b*) a voltage divider R_1 and R_2 to scale the input v_I, and (*c*) a comparator. The latter trips whenever v_I is such that $[R_1/(R_1 + R_2)]v_I = V_{REF}$. Denoting this special value of v_I as V_T, we get

$$V_T = (1 + R_2/R_1)V_{REF} \tag{9.5}$$

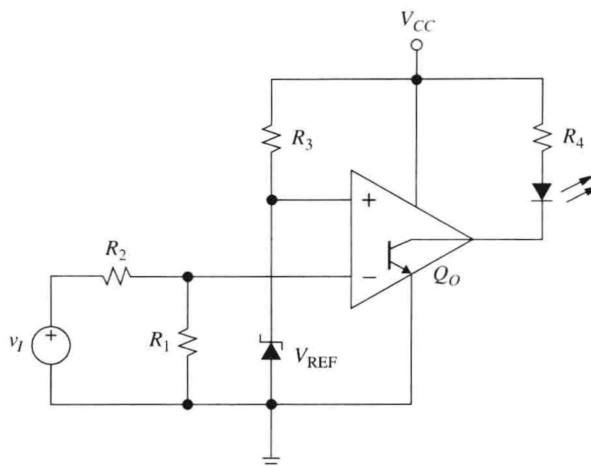


FIGURE 9.12
Basic level detector with optical indicator.

For $v_I < V_T$, Q_O is off and so is the LED. For $v_I > V_T$, Q_O saturates and makes the LED glow, thus providing an indication of when v_I rises above V_T. Interchanging the input pins will make the LED glow whenever v_I drops below V_T. The function of R_3 is to bias the reference diode, and that of R_4 is to set the LED current.

EXAMPLE 9.1. In the circuit of Fig. 9.12 let $V_{REF} = 2.0$ V, $R_1 = 20$ kΩ, and $R_2 = 30$ kΩ. Assuming a 339 comparator with $V_{OS} = 5$ mV (maximum) and $I_B = 250$ nA (maximum), estimate the worst-case error of your circuit.

Solution. In this circuit, I_P has no effect because $R_p \cong 0$. Since I_N flows out of the comparator, it raises the inverting-input voltage by $(R_1 \parallel R_2)I_N = 3$ mV (maximum) when the comparator is about to trip. The worst-case scenario occurs when V_{OS} adds in the same direction, for a net inverting-input rise of $V_{OS} + (R_1 \parallel R_2)I_N = 5 + 3 = 8$ mV (maximum). This has the same effect as an 8-mV drop in V_{REF}, giving $V_T = (1 + 30/20)(2 - 0.008) = 4.98$ V instead of $V_T = 5.00$ V.

If v_I is V_{CC} itself, the circuit will monitor its own power supply and function as an *overvoltage* indicator. If the input pins are interchanged with each other so that $v_N = V_{REF}$ and $v_P = v_I/(1 + R_2/R_1)$, then we have an *undervoltage* indicator.

EXAMPLE 9.2. Using comparators of the 339 type, an LM385 2.5-V reference diode ($I_R \cong 1$ mA), and two HLMP-4700 LEDs ($I_{LED} \cong 2$ mA and $V_{LED} \cong 1.8$ V), design a circuit that monitors a 12-V car battery and causes the first LED to glow whenever the battery voltage rises above 13 V, and causes the second LED to glow whenever it drops below 10 V.

Solution. We need two comparators, one for overvoltage and the other for undervoltage sensing. The comparators share the same reference diode, and in both cases v_I is the battery voltage V_{CC}. For the overvoltage circuit we need $13 = (1 + R_2/R_1)2.5$ and $R_4 = (13 - 1.8)/2$; use $R_1 = 10.0$ kΩ and $R_2 = 42.2$ kΩ, both 1%, and $R_4 = 5.6$ kΩ. For the undervoltage circuit we interchange the input pins, and we impose $10 = (1 + R_2/R_1)2.5$ and $R_4 = (10 - 1.8)/2$; use $R_1 = 10.0$ kΩ and $R_2 = 30.1$ kΩ, both 1%, and $R_4 = 3.9$ kΩ. To bias the reference diode, use $R_3 = (12 - 2.5)/1 \cong 10$ kΩ.

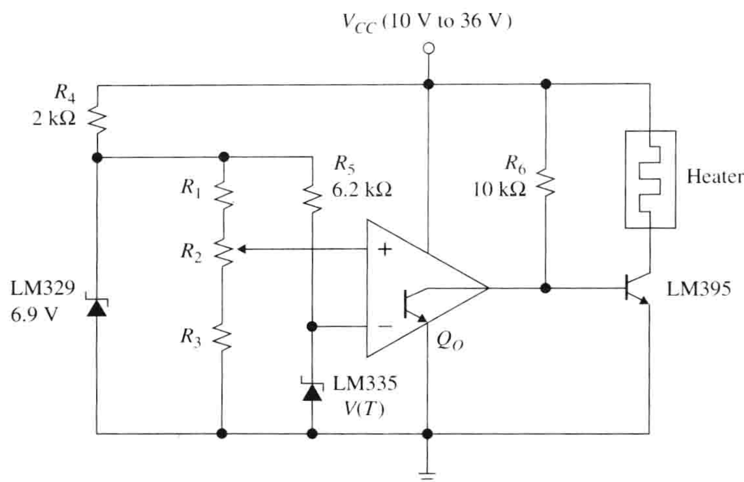


FIGURE 9.13
On-off temperature controller.

On-Off Control

Level detection can be applied to any physical variable that can be expressed in terms of a voltage via a suitable transducer. Typical examples are temperature, pressure, strain, position, fluidic level, and light or sound intensity. Moreover, the comparator can be used not only to monitor the variable, but also to control it.

Figure 9.13 shows a simple temperature controller, or thermostat. The comparator, a 339 type, uses the LM335 temperature sensor to monitor temperature, and the LM395 high-beta power transistor to switch a heater on and off in order to keep temperature at the setpoint established via R_2. The LM335 is an active reference diode designed to produce a temperature-dependent voltage according to $V(T) = T/100$, where T is absolute temperature, in kelvins. The purpose of R_5 is to bias the sensor. For the circuit to work over a wide range of supply voltages, the transducer-bridge voltage must be stabilized. This function is provided by the LM329 6.9-V reference diode, which is biased via R_4.

The circuit operates as follows. As long as temperature is above the setpoint, we have $v_N > v_P$; Q_O saturates and keeps the LM395-heater combination off. If, however, temperature drops below the setpoint, then $v_N < v_P$; Q_O is now in cutoff, thus diverting the current supplied by R_6 to the base of the LM395 transistor. The latter then saturates, turning the heater fully on.

Both the sensor and the heater are placed inside an oven and can be used, for instance, to thermostat a quartz crystal. This also forms the basis of *substrate thermostating,* a technique often used to stabilize the characteristics of voltage references and log/antilog amplifiers. We will see examples in Chapter 11 and 13.

EXAMPLE 9.3. In the circuit of Fig. 9.13 specify suitable resistances so that the set point can be adjusted anywhere between 50 °C and 100 °C by means of a 5-kΩ potentiometer.

Solution. Since $V(50\ °\text{C}) = (273.2 + 50)/100 = 3.232$ V, and $V(100\ °\text{C}) = 3.732$ V, the current through R_2 is $(3.732 - 3.232)/5 = 0.1$ mA. Consequently, $R_3 = 3.232/0.1 =$ 32.3 kΩ (use 32.4 kΩ, 1%), and $R_1 = (6.9 - 3.732)/0.1 = 31.7$ kΩ (use 31.6 kΩ, 1%).

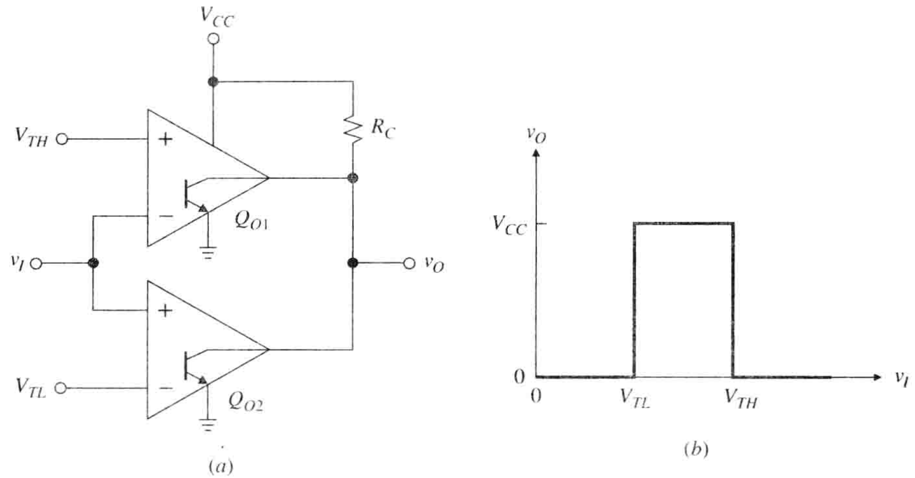


FIGURE 9.14
Window detector and its VTC.

Window Detectors

The function of a window detector, also called a *window comparator,* is to indicate when a given voltage falls within a specified *band,* or *window.* This function is implemented with a pair of level detectors, whose thresholds V_{TL} and V_{TH} define the lower and upper limits of the window. Referring to Fig. 9.14*a*, we observe that as long as $V_{TL} < v_I < V_{TH}$, both Q_{O1} and Q_{O2} are off, so R_C pulls v_O to V_{CC} to yield a high output. Should, however, v_I fall outside the range, the output BJT of one of the comparators will go on (Q_{O1} for $v_I > V_{TH}$, Q_{O2} for $v_I < V_{TL}$) and bring v_O near 0 V. Figure 9.14*b* shows the resulting VTC.

If R_C is replaced by an LED in series with a suitable current-limiting resistor, the LED will glow whenever v_I falls outside the window. If we wish the LED to glow whenever v_I falls inside the window, then we must insert an inverting stage between the comparators and the LED-resistor combination. An inverter example is offered by the 2N2222 BJT of Fig. 9.15.

The window detector shown monitors whether its own supply voltage is within tolerance. The top comparator pulls the base of the 2N2222 BJT low whenever V_{CC} drops below a given lower limit, and the bottom comparator pulls the base low whenever V_{CC} rises above a given upper limit; in either case the LED is off. For V_{CC} within tolerance, however, the output BJTs of both comparators are off, letting R_4 turn on the 2N2222 BJT and thus causing the LED to glow.

EXAMPLE 9.4. Specify suitable component values so that the LED of Fig. 9.15 glows for V_{CC} within the band 5 V $\pm 5\%$, which is the band usually required by digital circuits to work according to specification. Assume $V_{\text{LED}} \cong 1.5$ V, and impose $I_{\text{LED}} \cong 10$ mA and $I_{B(\text{2N2222})} \cong 1$ mA.

Solution. For $V_{CC} = 5 + 5\% = 5.25$ V we want $v_N = 2.5$ V for the bottom comparator; for $V_{CC} = 5 - 5\% = 4.75$ V we want $v_P = 2.5$ V for the top comparator. Using the

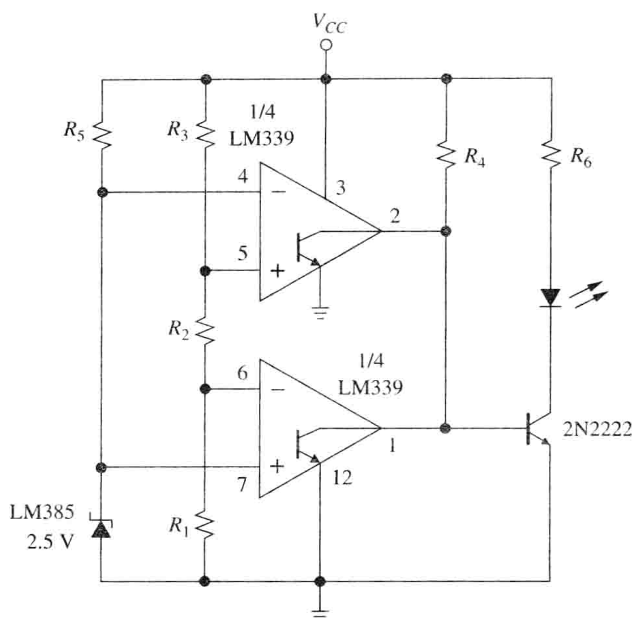


FIGURE 9.15
Power-supply monitor; LED glows as long as V_{CC} is within specification.

voltage divider formula twice gives $2.5/5.25 = R_1/(R_1 + R_2 + R_3)$, and $2.5/4.75 = (R_1 + R_2)/(R_1 + R_2 + R_3)$. Let $R_1 = 10.0$ kΩ; then we get $R_2 = 1.05$ kΩ and $R_3 = 10.0$ kΩ. Moreover, $R_4 = (5 - 0.7)/1 = 4.3$ kΩ, $R_5 = (5 - 2.5)/1 \cong 2.7$ kΩ, and $R_6 = (5 - 1.5)/10 \cong 330$ Ω.

Window comparators are used in production-line testing to sort out circuits that fail to meet a given tolerance. In this and other automatic test and measurement applications, V_{TL} and V_{TH} are usually provided by a computer via a pair of D-A converters.

Bar Graph Meters

A bar graph meter provides a visual indication of the input signal level. The circuit is a generalization of the window detector in that it partitions the input signal range into a string of consecutive windows, or steps, and uses a string of comparator-LED pairs to indicate the window within which the input falls at a given time. The larger the number of windows, the higher the resolution of the bar display.

Figure 9.16 shows the block diagram of the popular LM3914 bar graph meter. The upper and lower limits of the signal range are set by the user via the voltages applied to the reference low (R_{LO}) and the reference high (R_{HI}) input pins. An internal resistance string partitions this range into ten consecutive windows, and each comparator causes the corresponding LED to glow whenever v_I rises above the reference voltage available at the corresponding tap. The input level can be

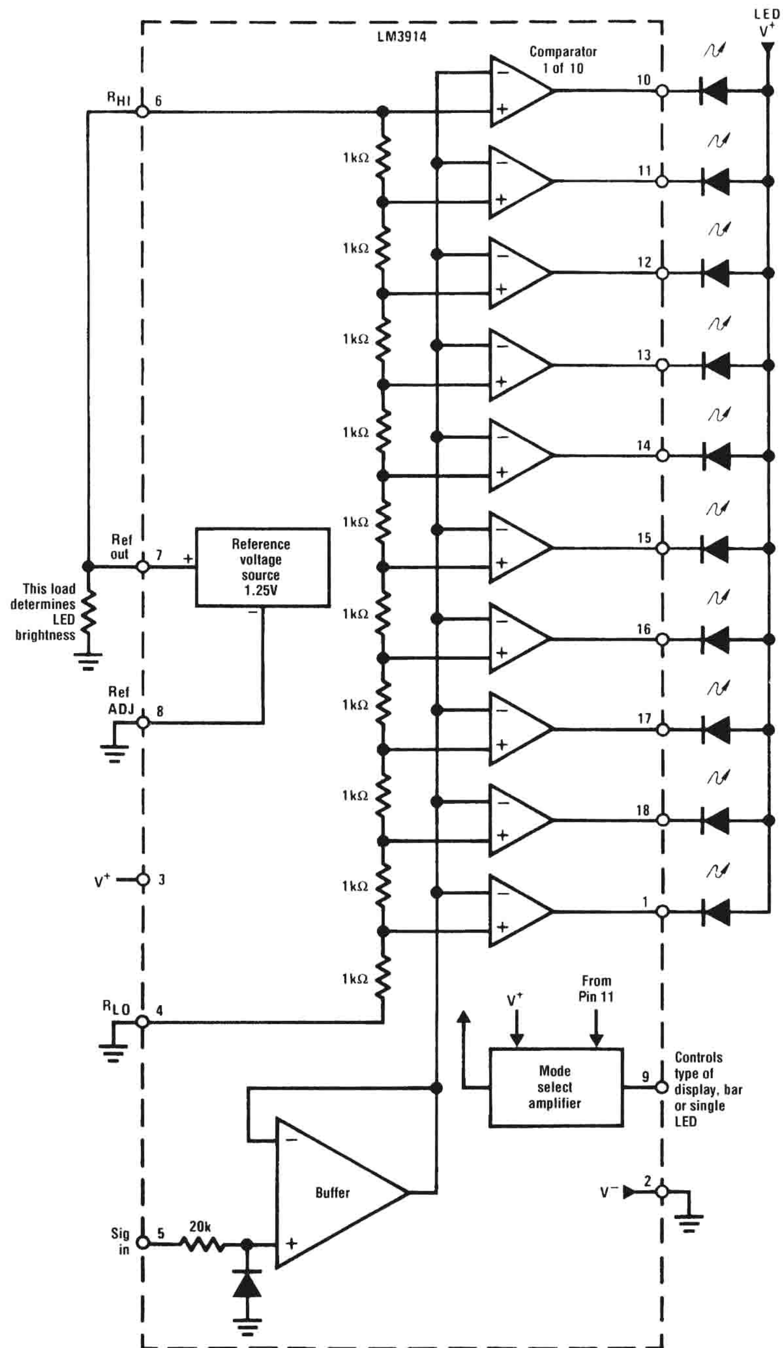


FIGURE 9.16
The LM3914 dot/bar display driver. (Courtesy of Texas Instruments.)

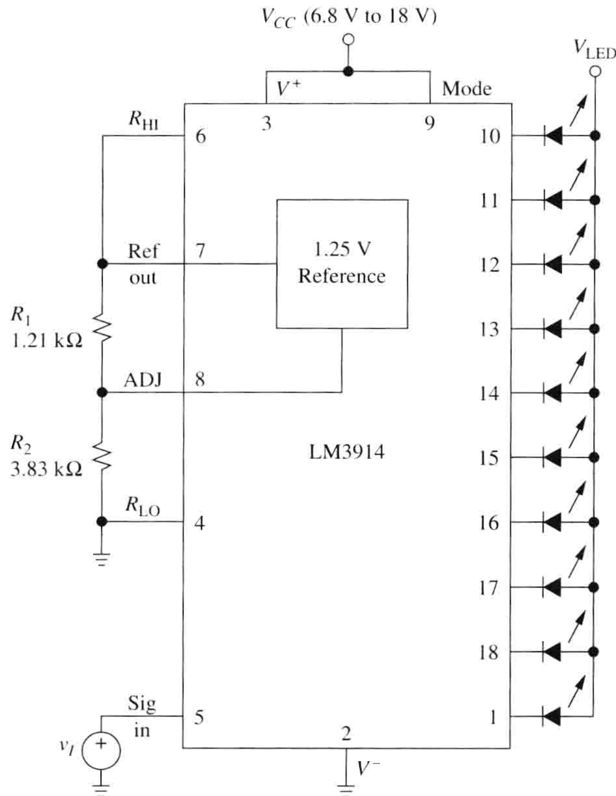


FIGURE 9.17
0-V to 5-V bar graph meter. (Courtesy of Texas Instruments.)

visualized either in bar graph form, or as a moving dot, depending on the logic level applied at the mode control pin 9.

The circuit also includes an input buffer to prevent loading the external source and a 1.25-V reference source to facilitate input range programming. With the connection of Fig. 9.16 the input range is from 0 V to 1.25 V; however, bootstrapping the reference source, as in Fig. 9.17, expands the upper limit to $(1 + R_2/R_1)1.25 + R_2 I_{ADJ}$, where I_{ADJ} is the current flowing out of pin 8. Since $I_{ADJ} \cong 75\ \mu A$, specifying R_2 in the low-kilohm range will make the $R_2 I_{ADJ}$ term negligible, so the input range is from 0 V to $(1 + R_2/R_1)1.25$ V. A variety of other configurations are possible, such as multiple-device cascading for greater resolution, and zero-center meter operation. Consult the data sheets for more details.

The LM3915 is similar to the LM3914, except that the resistance string values have been chosen to give 3-dB logarithmic steps. This type of display is intended for signals with wide dynamic ranges, such as audio level, power, and light intensity. The LM3916 is similiar to the LM3915, except that the steps are chosen to configure the device for VU meter readings, the type of readings commonly used in audio and radio applications.

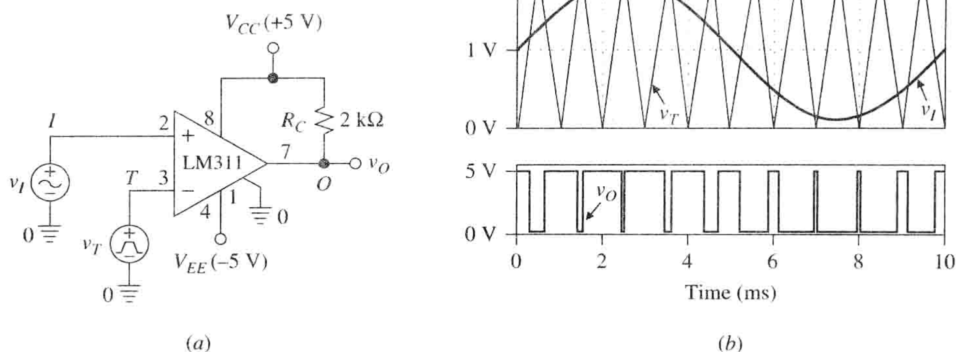


FIGURE 9.18
(*a*) Using the 311 macromodel to illustrate pulse-width modulation. (*b*) Waveforms.

Pulse-Width Modulation

If a voltage comparator is made to compare a slowly varying signal v_I against a high-frequency triangular wave v_T, the outcome v_O is a square wave with the same frequency as v_T, but with its symmetry controlled by v_I. This is illustrated in Fig. 9.18 for the case of a sinusoidal input v_I.

The degree of symmetry of v_O is expressed via the *duty cycle*

$$D(\%) = 100\frac{T_H}{T_L + T_H} \tag{9.6}$$

where T_L and T_H denote, respectively, the times spent by v_O in the low and the high state within a given cycle of v_{TR}. For instance, if v_O is high for 0.75 ms and low for 0.25 ms, then $D(\%) = 100 \times 0.75/(0.25 + 0.75) = 75\%$. It is readily seen that for the example illustrated we have

$$D(\%) = 100\frac{v_I}{2\text{ V}} \tag{9.7}$$

indicating that varying v_I over the range $0 < v_I < 2$ V varies D over the range $0\% < D < 100\%$. We can regard v_O as a train of pulses whose widths are controlled, or modulated, by v_I. *Pulse-width modulation* (PWM) finds application in signal transmission and power control.

9.3 SCHMITT TRIGGERS

Having investigated the behavior of high-gain amplifiers with no feedback, we now turn to amplifiers with positive feedback, also known as *Schmitt triggers.* While negative feedback tends to keep the amplifier within the linear region, positive feedback forces it into saturation. The two types of feedback are compared in Fig. 9.19. At power turn-on, both circuits start out with $v_O = 0$. However, any input disturbance that might try to force v_O away from zero will elicit opposite responses. The

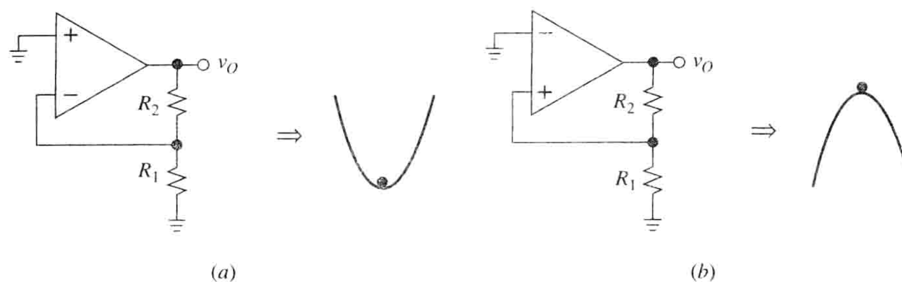


FIGURE 9.19
Mechanical models of (*a*) negative and (*b*) positive feedback.

amplifier with negative feedback will tend to neutralize the perturbation and return to the equilibrium state $v_O = 0$. Not so in the case of positive feedback, for now the reaction is in the same direction as the perturbation, indicating a tendency to reinforce rather than neutralize it. The ensuing regenerative effect will drive the amplifier into saturation, indicating two stable states, namely, $v_O = V_{OH}$ and $v_O = V_{OL}$.

In Fig. 9.19 negative feedback is likened to a ball at the bottom of a bowl, and positive feedback to a ball at the top of a dome. If we shake the bowl to simulate electronic noise, the ball will eventually return to its equilibrium position at the bottom, but shaking the dome will cause the ball to fall to either side.

Inverting Schmitt Trigger

The circuit of Fig. 9.20*a* uses a voltage divider to provide positive dc feedback around a 301 op amp. The circuit can be viewed as an inverting-type threshold detector whose threshold is controlled by the output. Since the output has two stable states, this threshold has two possible values, namely,

$$V_{TH} = \frac{R_1}{R_1 + R_2} V_{OH} \qquad V_{TL} = \frac{R_1}{R_1 + R_2} V_{OL} \tag{9.8}$$

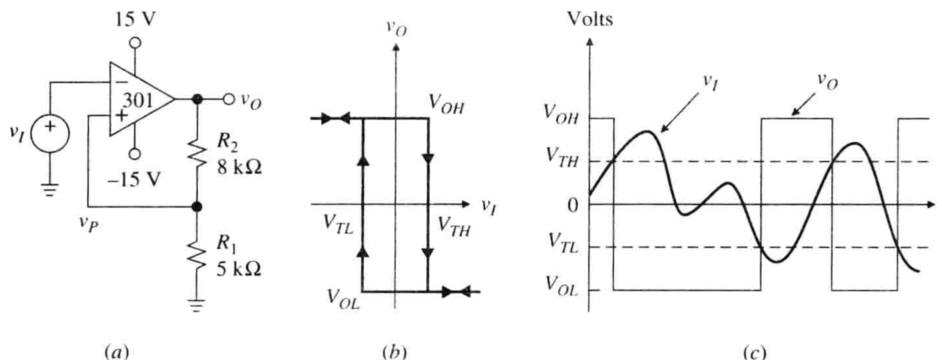


FIGURE 9.20
Inverting Schmitt trigger, VTC, and sample waveforms.

With the output saturating at ± 13 V, the component values shown give $V_{TH} = +5$ V and $V_{TL} = -5$ V, also expressed as $V_T = \pm 5$ V.

The best way to visualize circuit behavior is by deriving its VTC. Thus, for $v_I \ll 0$, the amplifier saturates at $V_{OH} = +13$ V, giving $v_P = V_{TH} = +5$ V. Increasing v_I moves the operating point along the upper segment of the curve until v_I reaches V_{TH}. At this juncture the regenerative action of positive feedback causes v_O to snap from V_{OH} to V_{OL} as fast as the amplifier can swing. This, in turn, causes v_P to snap from V_{TH} to V_{TL}, or from $+5$ V to -5 V. If we wish to change the output state again, we must now lower v_I all the way down to $v_P = V_{TL} = -5$ V, at which juncture v_O will snap back to V_{OH}. In summary, as soon as $v_N = v_I$ approaches $v_P = V_T$, v_O and, hence, v_P, snap away from v_N. This behavior is opposite to that of negative feedback, where v_N tracks v_P!

Looking at the VTC of Fig. 9.20*b*, we observe that when coming from the left, the threshold is V_{TH}, and when coming from the right it is V_{TL}. This can also be appreciated from the waveforms of Fig. 9.20*c*, where it is seen that during the times of increasing v_I the output snaps when v_I crosses V_{TH}, but during the times of decreasing v_I it snaps when v_I crosses V_{TL}. Note also that the horizontal portions of the VTC can be traveled in either direction, under external control, but the vertical portions can be traveled only *clockwise*, under the regenerative effect of positive feedback.

A VTC with two separate tripping points is said to exhibit *hysteresis*. The hysteresis *width* is defined as

$$\Delta V_T = V_{TH} - V_{TL} \tag{9.9}$$

and in the present case can be expressed as

$$\Delta V_T = \frac{R_1}{R_1 + R_2}(V_{OH} - V_{OL}) \tag{9.10}$$

With the component values shown, $\Delta V_T = 10$ V. If desired, ΔV_T can be varied by changing the ratio R_1/R_2. Decreasing this ratio will bring V_{TH} and V_{TL} closer together until, in the limit $R_1/R_2 \to 0$, the two vertical segments coalesce at the origin. The circuit is then an inverting zero-crossing detector.

Noninverting Schmitt Trigger

The circuit of Fig. 9.21*a* is similar to that of Fig. 9.20*a*, except that v_I is now applied at the noninverting side. For $v_I \ll 0$, the output will saturate at V_{OL}. If we want v_O to switch state, we must raise v_I to a high enough value to bring v_P to cross $v_N = 0$, since this is when the comparator trips. This value of v_I, aptly denoted as V_{TH}, must be such that $(V_{TH} - 0)/R_1 = (0 - V_{OL})/R_2$, or

$$V_{TH} = -\frac{R_1}{R_2}V_{OL} \tag{9.11a}$$

Once v_O has snapped to V_{OH}, v_I must be lowered if we want v_O to snap back to V_{OL}. The tripping voltage V_{TL} is such that $(V_{OH} - 0)/R_2 = (0 - V_{TL})/R_1$, or

$$V_{TL} = -\frac{R_1}{R_2}V_{OH} \tag{9.11b}$$

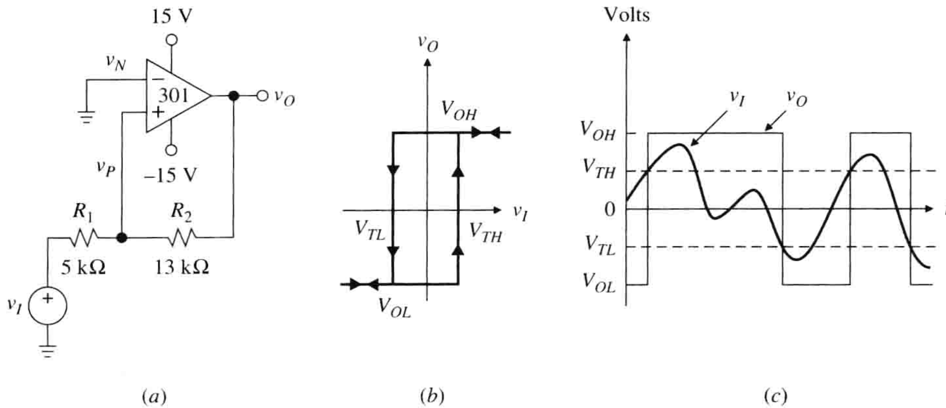


FIGURE 9.21
Noninverting Schmitt trigger, VTC, and sample waveforms.

The resulting VTC, shown in Fig. 9.21*b*, differs from that of Fig. 9.20*b* in that the vertical segments are traveled in the *counterclockwise* direction. The output waveform is similar to that of the inverting Schmitt trigger, except for a reversal in polarity. The hysteresis width is now

$$\Delta V_T = \frac{R_1}{R_2}(V_{OH} - V_{OL}) \tag{9.12}$$

and it can be varied by changing the ratio R_1/R_2. In the limit $R_1/R_2 \to 0$ we obtain a noninverting zero-crossing detector.

VTC Offsetting

In single-supply Schmitt triggers the need arises to offset the VTC so that it lies entirely within the first quadrant. The circuit of Fig. 9.22*a* achieves the positive

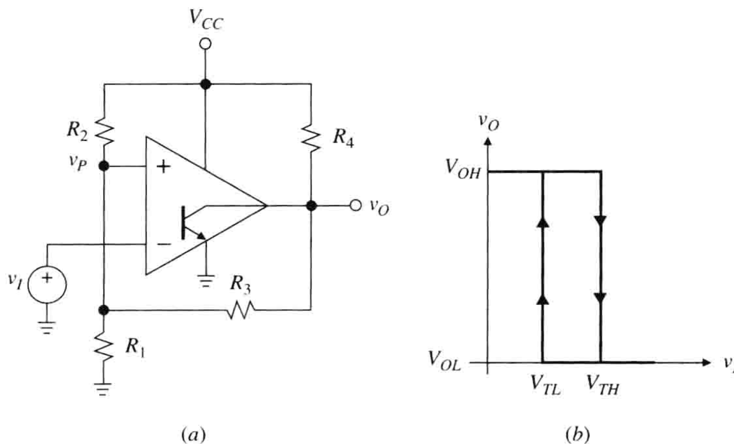


FIGURE 9.22
Single-supply inverting Schmitt trigger.

offset depicted in Fig. 9.22*b* by using a pullup resistance R_2. To find suitable design equations, we apply the superposition principle and write

$$v_P = \frac{R_1 \parallel R_3}{(R_1 \parallel R_3) + R_2} V_{CC} + \frac{R_1 \parallel R_2}{(R_1 \parallel R_2) + R_3} v_O$$

As we know, the circuit gives $V_{OL} \cong 0$ V. To achieve $V_{OH} \cong V_{CC}$, we specify $R_4 \ll R_3 + (R_1 \parallel R_2)$. Then, imposing $v_P = V_{TL}$ for $v_O = V_{OL} = 0$, and $v_P = V_{TH}$ for $v_O = V_{OH} = V_{CC}$, we get

$$V_{TL} = \frac{R_1 \parallel R_3}{(R_1 \parallel R_3) + R_2} V_{CC} \qquad V_{TH} = \frac{R_1}{R_1 + (R_2 \parallel R_3)} V_{CC}$$

Rearranging gives

$$\frac{1}{R_2} = \frac{V_{TL}}{V_{CC} - V_{TL}} \left(\frac{1}{R_1} + \frac{1}{R_3} \right) \qquad \frac{1}{R_1} = \frac{V_{CC} - V_{TH}}{V_{TH}} \left(\frac{1}{R_2} + \frac{1}{R_3} \right) \tag{9.13}$$

Since we have two equations and four unknown resistances, we fix two, say, R_4 and $R_3 \gg R_4$, and then solve for the other two.

EXAMPLE 9.5. Let the comparator of Fig. 9.22*a* be the LM339 type with $V_{CC} = 5$ V. Specify suitable resistances for $V_{OL} = 0$ V, $V_{OH} = 5$ V, $V_{TL} = 1.5$ V, and $V_{TH} = 2.5$ V.

Solution. Let $R_4 = 2.2$ kΩ (a reasonable value) and let $R_3 = 100$ kΩ (which is much greater than 2.2 kΩ). Then, $1/R_2 = (1.5/3.5)(1/R_1 + 1/100)$ and $1/R_1 = 1/R_2 + 1/100$. Solving yields $R_1 = 40$ kΩ (use 39 kΩ) and $R_2 = 66.7$ kΩ (use 68 kΩ).

Figure 9.23*a* shows the noninverting realization of the single-supply Schmitt trigger. Here, the function of R_1 and R_2 is to provide a suitable bias for v_N. Imposing $R_5 \ll R_3 + R_4$ to ensure $V_{OH} \cong V_{CC}$, and following a similar line of reasoning, one can readily show (see Problem 9.10) that

$$\frac{R_3}{R_4} = \frac{V_{TH} - V_{TL}}{V_{CC}} \qquad \frac{R_2}{R_1} = \frac{V_{CC} - V_{TL}}{V_{TH}} \tag{9.14}$$

These equations are used to achieve the desired V_{TL} and V_{TH}.

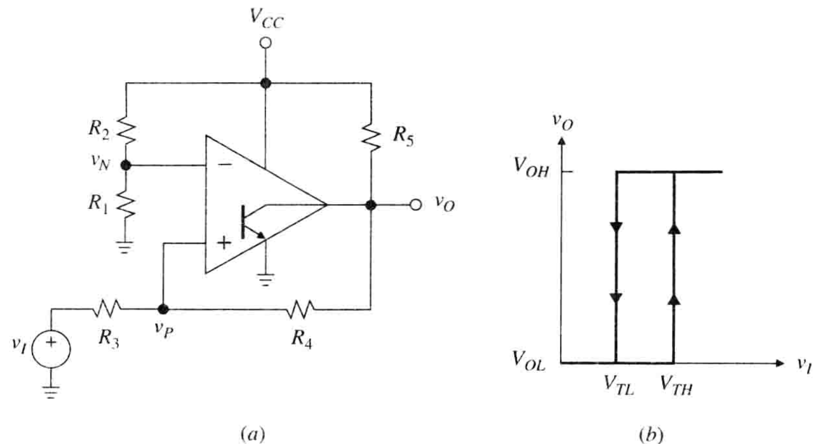


FIGURE 9.23
Single-supply noninverting Schmitt trigger.

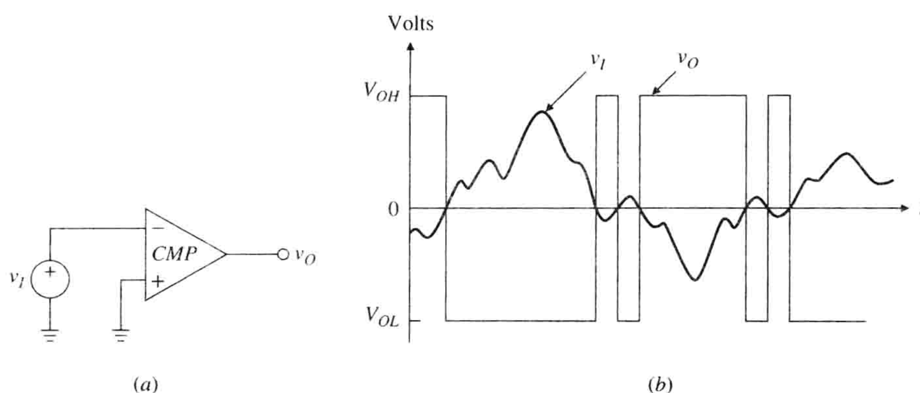


FIGURE 9.24
Comparator chatter.

Eliminating Comparator Chatter

When processing slowly varying signals, comparators tend to produce multiple output transitions, or bounces, as the input crosses the threshold region. Figure 9.24 shows an example. Referred to as *comparator chatter*, these bounces are due to ac noise invariably superimposed on the input signal, especially in industrial environments. As this signal crosses the threshold region, noise is amplified with the full open-loop gain, causing output chatter. For instance, the LM311 comparator, whose gain is typically 200 V/mV, requires an input noise spike of only $(5/200{,}000) = 25\ \mu\text{V}$ to cause a 5-V output swing. Chatter is unacceptable in counter-based applications.

The problem is eliminated with the help of hysteresis, as shown in Fig. 9.25. In this case, as soon as v_I crosses the present threshold, the circuit snaps and activates the other threshold, so v_I must swing back to the new threshold in order to make v_O snap again. Making the hysteresis width greater than the maximum peak-to-peak amplitude of noise prevents spurious output transitions.

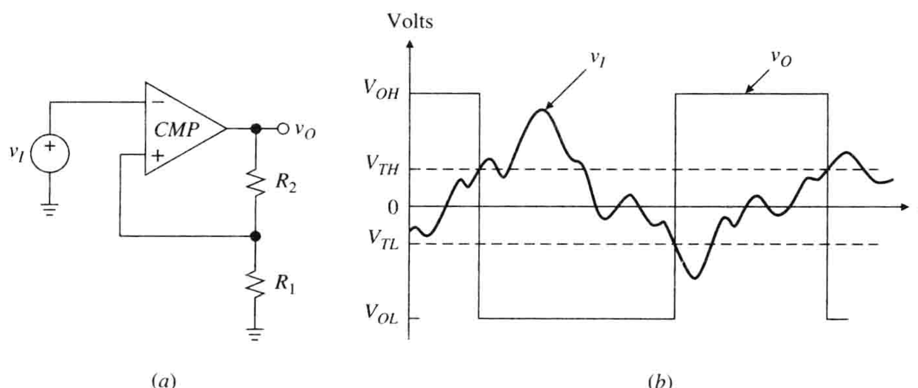


FIGURE 9.25
Using hysteresis to eliminate chatter.

Even in situations where the input signal is relatively clean, it always pays to introduce a small amount of hysteresis, say, a few millivolts, to stave off potential oscillations due to stray ac feedback caused by parasitic capacitances and the distributed impedances of the power-supply and ground busses. This stabilization technique is particularly important in flash A-D converters.

Hysteresis in On-Off Controllers

Hysteresis is used in on-off control to avoid overfrequent cycling of pumps, furnaces, and motors. Consider, for instance, the temperature controller discussed in connection with Fig. 9.13. We can easily turn it into a home thermostat by having the comparator drive a power switch like a relay or a triac to turn a home furnace on or off. Starting with temperatures below the setpoint, the comparator will activate the furnace and cause the temperature to rise. This rise is monitored by the temperature sensor and conveyed to the comparator in the form of an increasing voltage. As soon as the temperature reaches the setpoint, the comparator will trip and shut off the furnace. However, the smallest temperature drop following furnace shut off will suffice to trip the comparator back to the active state. As a result, the furnace will be cycled on and off at a rapid pace, a very taxing affair.

In general, temperature need not be regulated to such a sharp degree. Allowing a hysteresis of a few degrees will still ensure a comfortable environment and yet reduce furnace cycling significantly. This we achieve by providing a small amount of hysteresis.

EXAMPLE 9.6. Modify the temperature controller of Example 9.3 to ensure a hysteresis of about $\pm 1\ ^\circ\text{C}$. The LM395 power BJT has typically $V_{BE(\text{on})} = 0.9$ V.

Solution. Connect a positive-feedback resistance R_F between the output v_O and the noninverting input v_P of the comparator, so that $\Delta v_P = \Delta v_O R_W/(R_W + R_F)$, where R_W is the equivalent resistance presented to R_F by the wiper. With the wiper in the middle, $R_W = (R_1 + R_2/2)\|(R_3 + R_2/2) = 17.2\ \text{k}\Omega$. Using $\Delta v_O = 0.9$ V and $\Delta v_P = \pm 1 \times 10\ \text{mV} = 20$ mV, and solving, we get $R_F \cong 750\ \text{k}\Omega$.

9.4 PRECISION RECTIFIERS

A *half-wave rectifier* (HWR) is a circuit that passes only the positive (or only the negative) portion of a wave, while blocking out the other portion. The transfer characteristic of the positive HWR, pictured in Fig. 9.26*a*, is

$$v_O = v_I \qquad \text{for } v_I > 0 \tag{9.15a}$$

$$v_O = 0 \qquad \text{for } v_I < 0 \tag{9.15b}$$

A *full-wave rectifier* (FWR), besides passing the positive portion, inverts and then passes also the negative portion. Its transfer characteristic, depicted in Fig. 9.26*b*, is $v_O = v_I$ for $v_I > 0$, and $v_O = -v_I$ for $v_I < 0$, or, more concisely,

$$v_O = |v_I| \tag{9.16}$$

An FWR is also referred to as an *absolute-value circuit.*

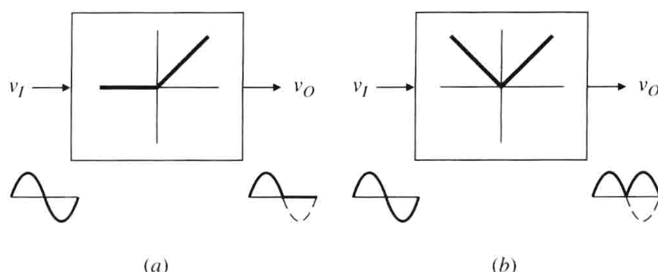


FIGURE 9.26
Half-wave rectifier (HWR) and full-wave rectifier (FWR).

Rectifiers are implemented using nonlinear devices such as diodes. The nonzero forward-voltage drop $V_{D(\text{on})}$ of a practical diode may cause intolerable errors in low-level signal rectification. As we shall see, this shortcoming is avoided by placing the diode inside the negative-feedback path of an op amp.

Half-Wave Rectifiers

The analysis of the circuit of Fig. 9.27 is facilitated if we consider the cases $v_I > 0$ and $v_I < 0$ separately.

1. $v_I > 0$: In response to a positive input, the op amp output v_{OA} will also swing positive, turning on the diode and thus creating the negative-feedback path shown in Fig. 9.28*a*. This allows us to apply the virtual-short principle and write $v_O = v_I$. We observe that to make v_O track v_I, the op amp rides its output a diode drop above v_O, that is, $v_{OA} = v_O + V_{D(\text{on})} \cong v_O + 0.7$ V. Placing the diode within the feedback loop in effect eliminates any errors due to its forward-voltage drop. To emphasize this dramatic effect of negative feedback, the diode-op amp combination is referred to as a *superdiode.*
2. $v_I < 0$: Now the op amp output swings negative, turning the diode off and thus causing the current through R to go to zero. Hence, $v_O = 0$. As pictured in Fig. 9.28*b*, the op amp is now operating in the open-loop mode, and since $v_P < v_N$, the output saturates at $v_{OA} = V_{OL}$. With $V_{EE} = -15$ V, $v_{OA} \cong -13$ V.

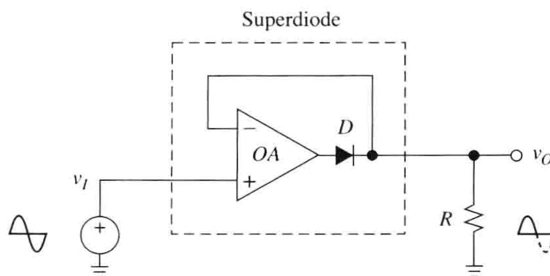


FIGURE 9.27
Basic half-wave rectifier.

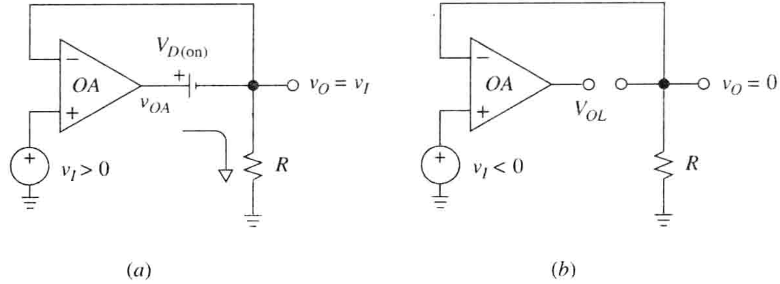


FIGURE 9.28
Equivalent circuits of the basic HWR for (*a*) positive and (*b*) negative inputs.

A disadvantage of this circuit is that when v_I changes from negative to positive, the op amp output has to come out of saturation and then swing all the way from $v_{OA} = V_{OL} \cong -13$ V to $v_{OA} \cong v_I + 0.7$ V in order to close the feedback loop. All this takes time, and if v_I has changed appreciably meanwhile, v_O may exhibit intolerable distortion. The improved HWR of Fig. 9.29*a* alleviates this inconvenience by using a second diode to clamp the negative saturation level just a diode drop below ground. Proceeding as usual, we identify two cases:

1. $v_I > 0$: A positive input causes D_1 to conduct, thus creating a negative-feedback path around the op amp. By the virtual-ground principle we have $v_N = 0$, indicating that D_1 now clamps the op amp output at $v_{OA} = -V_{D1(\text{on})}$. Moreover, D_2 is off, so no current flows through R_2 and, hence, $v_O = 0$.
2. $v_I < 0$: A negative input causes the op amp output to swing positive, thus turning D_2 on. This creates an alternative negative-feedback path via D_2 and R_2, which still ensures $v_N = 0$. Clearly, D_1 is now off, so the current sourced by the op amp to R_2 must equal the current sunk by v_I from R_1, or $(v_O - 0)/R_2 = (0 - v_I)/R_1$. This gives $v_O = (-R_2/R_1)v_I$. Moreover, $v_{OA} = v_O + V_{D2(\text{on})}$.

Circuit behavior is summarized as

$$v_O = 0 \qquad \text{for } v_I > 0 \tag{9.17a}$$

$$v_O = -(R_2/R_1)v_I \qquad \text{for } v_I < 0 \tag{9.17b}$$

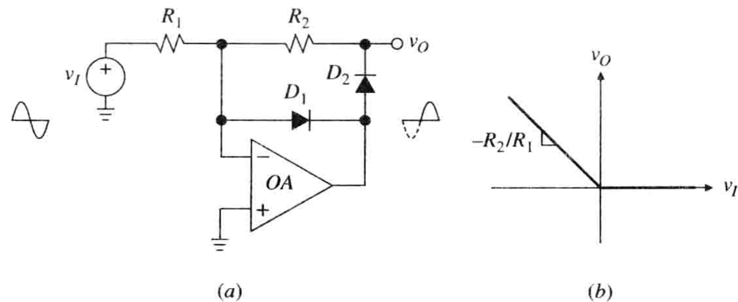


FIGURE 9.29
Improved HWR and its VTC.

and the VTC is shown in Fig. 9.29*b*. In words, the circuit acts as an inverting HWR with gain. The op amp output v_{OA} still rides a diode drop above v_O when $v_O > 0$; however, when $v_O = 0$, v_{OA} is clamped at about -0.7 V, that is, within the linear region. Consequently, the absence of saturation-related delays and the reduced output voltage swing result in much improved dynamics.

Full-Wave Rectifiers

One way of synthesizing the absolute value of a signal is by combining the signal itself with its inverted half-wave rectified version in a 1-to-2 ratio, as shown in Fig. 9.30. Here OA_1 provides inverting half-wave rectification, and OA_2 sums v_I and the HWR output v_{HW} in a 1-to-2 ratio to give $v_O = -(R_5/R_4)v_I - (R_5/R_3)v_{HW}$. Considering that $v_{HW} = -(R_2/R_1)v_I$ for $v_I > 0$, and $v_{HW} = 0$ for $v_I < 0$, we can write

$$v_O = A_p v_I \qquad \text{for } v_I > 0\text{ V} \tag{9.18a}$$

$$v_O = -A_n v_I \qquad \text{for } v_I < 0\text{ V} \tag{9.18b}$$

where

$$A_n = \frac{R_5}{R_4} \qquad A_p = \frac{R_2 R_5}{R_1 R_3} - A_n \tag{9.19}$$

We want both halves of the input wave to be amplified by the same gain $A_p = A_n = A$, for then we can write $v_O = Av_I$ for $v_I > 0$ and $v_O = -Av_I$ for $v_I < 0$, or, concisely,

$$v_O = A|v_I| \tag{9.20}$$

One way of achieving this goal is by imposing $R_1 = R_2 = R_4 = R$, $R_3 = R/2$, and $R_5 = AR$, as shown; then, $A = R_5/R$.

Because of resistance tolerances, A_p and A_n will generally differ from each other. Their difference

$$A_p - A_n = \frac{R_2 R_5}{R_1 R_3} - 2\frac{R_5}{R_4}$$

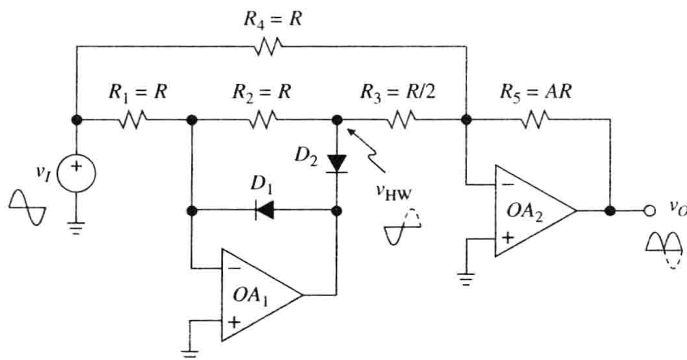


FIGURE 9.30
Precision FWR, or absolute-value circuit.

is maximized when R_2 and R_4 are maximized and R_1 and R_3 are minimized. (R_5 can be ignored since it appears in both terms.) Denoting percentage tolerance as p and substituting $R_2 = R_4 = R(1+p)$ and $R_1 = 2R_3 = R(1-p)$ gives

$$|A_p - A_n|_{\max} = 2A\left(\frac{1+p}{(1-p)^2} - \frac{1}{1+p}\right)$$

where $A = R_5/R$. For $p \ll 1$ we can ignore higher-order powers of p and use the approximations $(1 \pm p)^{-1} \cong (1 \mp p)$. This allows us to estimate the *maximum percentage difference* between A_p and A_n as

$$100\left|\frac{A_p - A_n}{A}\right|_{\max} \cong 800p$$

For instance, with 1% resistances, A_p and A_n may differ from each other by as much as about $800 \times 0.01 = 8\%$. To minimize this error, we can either use more precise resistors, such as laser-trimmed IC resistor arrays, or trim one of the first four resistors, say, R_2.

The alternative FWR realization of Fig. 9.31 requires only two matched resistors. For $v_I > 0$, D_1 is on, allowing OA_1 to keep its inverting input at virtual ground. With the output of OA_1 clamped at $-V_{D_1(\text{on})}$, D_2 is off, allowing R_4 to transmit v_I to OA_2. The latter, acting as a noninverting amplifier, gives $v_O = A_p v_I$,

$$A_p = 1 + \frac{R_3}{R_2}$$

For $v_I < 0$, D_1 is off and D_2 is forward biased by R_4. OA_1 still keeps its inverting input at virtual ground, but via the feedback path D_2-OA_2-R_3-R_2. By KCL, $(0 - v_I)/R_1 = (v_O - 0)/(R_2 + R_3)$, or $v_O = -A_n v_I$,

$$A_n = \frac{R_2 + R_3}{R_1}$$

Imposing $A_p = A_n = A$ allows us to write concisely $v_O = A|v_I|$. This condition is met by imposing $R_1 = R_2 = R$ and $R_3 = (A-1)R$, as shown. Clearly, only two matched resistances are needed.

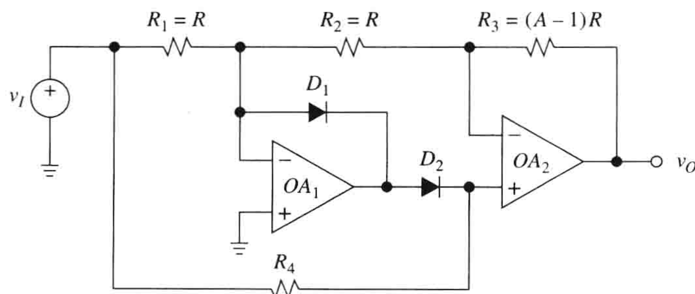


FIGURE 9.31
FWR using only two matched resistors.

Ac-dc Converters

The most common application of precision absolute-value circuits is ac-dc conversion, that is, the generation of a dc voltage proportional to the amplitude of a given ac wave. To accomplish this task, the ac signal is first full-wave rectified, and then low-pass filtered to synthesize a dc voltage. This voltage is the *average* of the rectified wave,

$$V_{avg} = \frac{1}{T}\int_0^T |v(t)|\, dt$$

where $v(t)$ is the ac wave and T is its period. Substituting $v(t) = V_m \sin(2\pi ft)$, where V_m is the peak amplitude and $f = 1/T$ is the frequency, gives

$$V_{avg} = (2/\pi)V_m = 0.637V_m$$

An ac-dc converter is calibrated so that when fed with an ac signal it gives its *root-mean-square (rms)* value,

$$V_{rms} = \left(\frac{1}{T}\int_0^T v^2(t)\, dt\right)^{1/2}$$

Substituting $v(t) = V_m \sin(2\pi ft)$ and integrating gives

$$V_{rms} = V_m/\sqrt{2} = 0.707V_m$$

The relationships between average and rms values and peak value are depicted in Fig. 9.32*a*. These relationships, which hold for sinusoidal waves but not necessarily for other waveforms, indicate that in order to obtain V_{rms} from V_{avg}, we need to multiply the latter by $(1/\sqrt{2})/(2/\pi) = 1.11$. The complete block diagram of an ac-dc converter is thus as in Fig. 9.32*b*.

Figure 9.33 shows a practical ac-dc converter implementation. The gain of 1.11 V/V is adjusted by means of the 50-kΩ pot, and the capacitance provides low-pass filtering with cutoff frequency $f_0 = 1/2\pi R_5 C$, where R_5 is the net resistance in parallel with C, or $1.11 \times 200 = 222$ kΩ. Hence, $f_0 = 0.717$ Hz. Using the LT1122 fast-settling JFET-input op amps allows the circuit to process a 10-V peak-to-peak ac signal with a 2-MHz bandwidth.

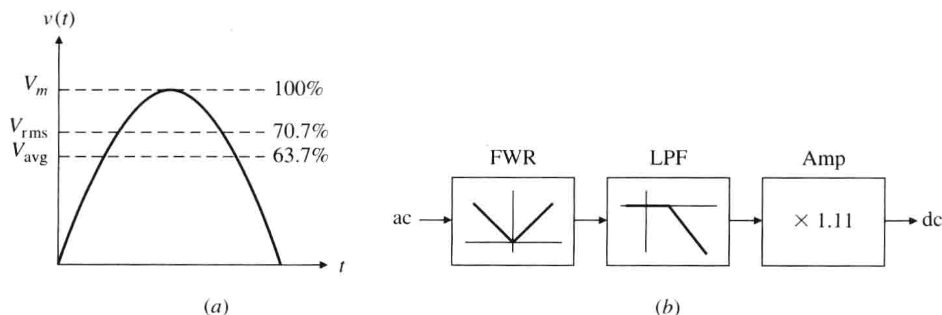


FIGURE 9.32
(*a*) Relationship between V_{rms} and V_m, and between V_{avg} and V_m. (*b*) Block diagram of an ac-dc converter.

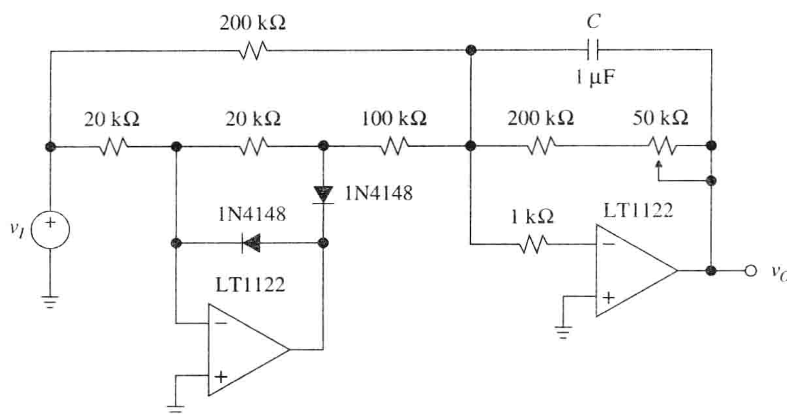


FIGURE 9.33
Wideband ac-dc converter.

The capacitance must be sufficiently large to keep the residual output ripple within specified limits. This requires that f_0 be well below the minimum operating frequency $f_{\min}$. Since an FWR doubles the frequency, the criterion for specifying C becomes

$$C \gg \frac{1}{4\pi R_5 f_{\min}}$$

As a conservative rule of thumb, C should exceed the right-hand term by the *inverse* of the *fractional ripple error* that can be tolerated at the output. For instance, for a 1% ripple error, C should be about $1/0.01 = 100$ times as large as the right-hand term. To remain within this error all the way down to the low end of the audio range, so that $f_{\min} = 20$ Hz, the above circuit would require $C = 100/(4\pi \times 222 \times 10^3 \times 20) \cong 1.8\ \mu$F.

9.5 ANALOG SWITCHES

Many circuits require electronic switches, that is, switches whose state is voltage-programmable. Chopper amplifiers, D-A converters, function generators, S/H amplifiers, and switching power supplies are common examples. Switches are also used to route signals in data acquisition systems, and to reconfigure circuits in programmable instrumentation.

As depicted in Fig. 9.34*a*, SW closes or opens, depending on the logic level at the control input C/O. When SW is closed, it drops zero voltage regardless of the current, and when SW is open, it draws zero current regardless of the voltage, thus giving the characteristic of Fig. 9.34*b*. This behavior can be approximated by any device capable of high on-off resistance ratios, such as field-effect transistors (FETs). An FET acts as a variable resistor called *channel,* whose resistance is controlled by the voltage applied between a control terminal called *gate* G and one of the channel terminals. These terminals, called *source* S and *drain* D, are usually interchangeable because the FET structure is symmetric.

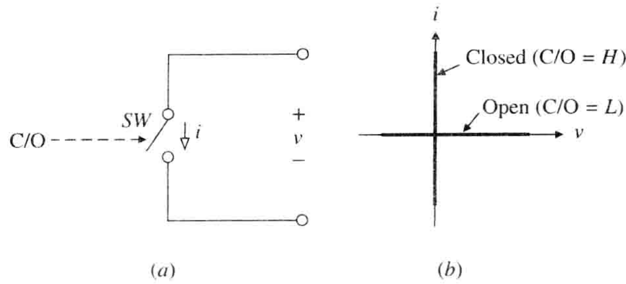


FIGURE 9.34
Ideal switch and its i-v characteristics.

JFET Switches

Figure 9.35 shows the characteristics of the n-channel JFET, or n-JFET for short. Each curve represents the i-v characteristic of the channel for a different value of the control voltage v_{GS} applied between gate and source. For $v_{GS} = 0$ the channel is highly conductive, this being the reason why JFETs are said to be *normally on* devices. Making v_{GS} progressively more negative reduces channel conductivity until a cutoff threshold $V_{GS(\text{off})} < 0$ is reached, such that for $v_{GS} \leq V_{GS(\text{off})}$ conductivity drops to zero and the channel acts as an open circuit. $V_{GS(\text{off})}$ is typically in the range of -0.5 V to -10 V, depending on the device.

In switch applications we are interested only in two curves, the ones corresponding to $v_{GS} = 0$ and $v_{GS} \leq V_{GS(\text{off})}$. The former is highly nonlinear; however, when the channel is used as a closed switch, its operation is near $v_{DS} = 0$ V, where the curve is fairly steep and linear. The slope is inversely proportional to a resistance $r_{ds(\text{on})}$ called the *dynamic resistance* of the channel,

$$\frac{di_D}{dv_{DS}} = \frac{1}{r_{ds(\text{on})}} \tag{9.21}$$

For ideal switch operation, this resistance should be zero; in practice, it is typically in the range of $10^2\ \Omega$ or less, depending on the device type.

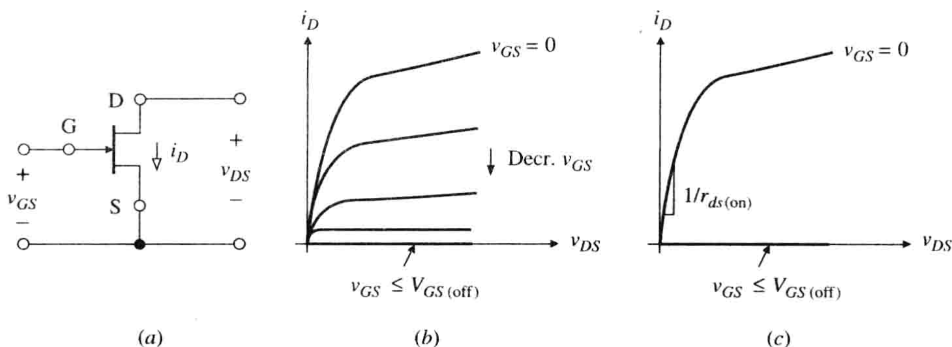


FIGURE 9.35
The n-channel JFET ($V_{GS(\text{off})} < 0$), and its i-v characteristics.

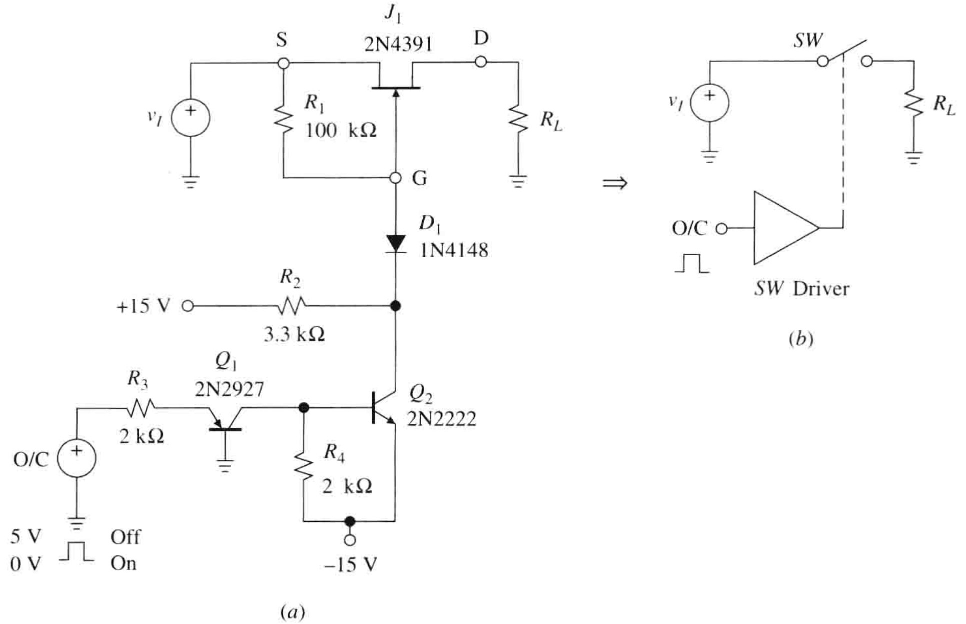


FIGURE 9.36
The n-channel JFET as a switch.

When the channel is off, its resistance is virtually infinite. The only currents of potential concern in this case are the leakage currents, namely, the *drain cutoff current* $I_{D(\text{off})}$ and the *gate reverse current* I_{GSS}. At ambient temperature these currents are typically in the picoampere range; however, they double with about every 10 °C increase. This can be of concern in certain applications, as we shall see.

A popular n-JFET switch is the 2N4391 (Siliconix), whose room-temperature ratings are: $-4\text{ V} \le V_{GS(\text{off})} \le -10$ V, $r_{ds(\text{on})} \le 30\ \Omega$, $I_{D(\text{off})} \le 100$ pA, $I_{GSS} \le 100$ pA flowing out of the gate, turn-on delay ≤ 15 ns, and turn-off delay ≤ 20 ns. Figure 9.36 illustrates a typical switch application. The function of the switch is to provide a make/break connection between a source v_I and a load R_L, whereas the function of the switch driver is to translate the TTL-compatible logic command O/C to the proper gate drive.

With O/C low ($\cong 0$ V), the E-B junction of Q_1 is off, so both Q_1 and Q_2 are off. By the pullup action of R_2, D_1 is reverse-biased, allowing R_1 to keep the gate at the same potential as the channel. We thus have $v_{GS} = 0$ regardless of v_I, so the switch is heavily on.

With O/C high ($\cong 5$ V), Q_1 conducts and forces Q_2 to saturate, thus pulling the gate close to -15 V. With a gate voltage this negative, the switch is off. To prevent J_1 from inadvertently going on, we must limit v_I in the negative direction,

$$v_{I(\text{min})} = V_{EE} + V_{CE2(\text{sat})} + V_{D1(\text{on})} - V_{GS(\text{off})} \tag{9.22}$$

For instance, with $V_{GS(\text{off})} = -4$ V we obtain $v_{I(\text{min})} \cong -15+0.1+0.7-(-4) \cong -10$ V, indicating that the circuit will operate properly only as long as the input is above -10 V.

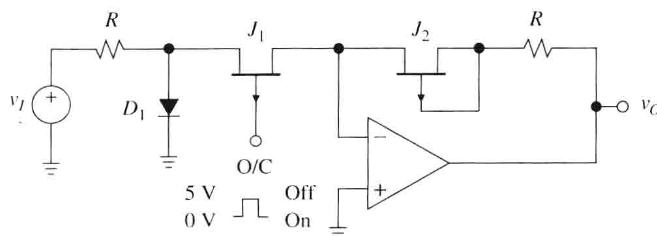


FIGURE 9.37
Analog ground switch using p-channel JFETs.

For high-speed operation,[2] connect a 100-pF capacitor between the control input and the base of Q_2 to speed up the turn-on and turn-off times of Q_2, and an HP2810 Schottky diode between the base and collector of Q_2 (anode at the base) to eliminate the storage delay of Q_2. JFET drivers as well as JFET-driver combinations are available in IC form from a variety of manufacturers.

The configuration of Fig. 9.36 requires a dedicated driver because the switch must ride with the signal v_I. If the switch is allowed to remain at a nearly constant potential, such as the virtual-ground potential of an op amp, then the driver can be simplified or even eliminated, as the configuration of Fig. 9.37 shows. Referred to as *analog ground switch* or *current switch,* this configuration uses a p-JFET designed for direct compatibility with standard logic levels. The p-JFET is similar to the n-JFET, except that the cutoff voltage is now positive, or $V_{GS(\text{off})} > 0$. Moreover, the fabrication of p-JFETs is compatible with low-cost bipolar technology. The switch operates as follows.

When the control input O/C is low, we have $v_{GS1} \cong 0$, indicating that J_1 is heavily on. To compensate for the presence of $r_{ds1(\text{on})}$, a dummy JFET J_2 is used in the feedback path of the op amp with the gate and source tied together to keep it permanently on. J_1 and J_2 are matched devices to ensure $r_{ds2(\text{on})} = r_{ds1(\text{on})}$ and, hence, $v_O/v_I = -1$ V/V.

When O/C is high, or $v_{GS1} > V_{GS1(\text{off})}$, J_1 is off and signal propagation is thus inhibited, so now $v_O/v_I = 0$. D_1 provides a clamping function to prevent the channel from inadvertently turning on during the positive alternations of v_I. Summarizing, the circuit provides unity gain when O/C is low, and zero gain when O/C is high.

The principle of Fig. 9.37 is especially useful in summing-amplifier applications. Replicating the input resistor-diode-switch combination k times gives a *k-channel analog multiplexer,* a device widely used in data acquisition and audio signal switching. The AH5010 quad switch consists of four p-FET switches and relative diode clamps plus a dummy FET in the same package. With an external op amp and five resistors, one can implement a four-channel multiplexer, and by cascading multiple AH5010s, one can expand to virtually any number of channels.

MOSFET Switches

Since MOS technology forms the basis of digital VLSI, MOS switches are particularly attractive when analog and digital functions must coexist on the same chip. MOSFETs are available both in normally on, or *depletion,* versions, and in normally

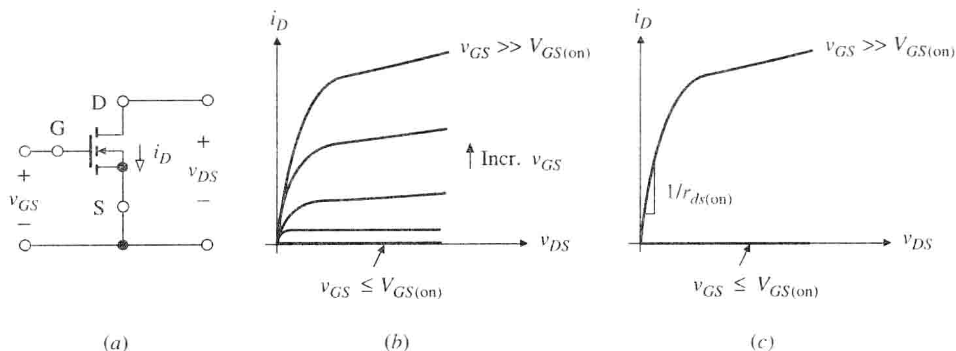


FIGURE 9.38
The enhancement n-channel MOSFET ($V_{GS(on)} > 0$), and its i-v characteristics.

off, or *enhancement,* versions. The latter are by far the most common, since they form the basis of CMOS technology.

Figure 9.38 shows the characteristics of the enhancement n-channel MOSFET, or n-MOSFET, for short. Its behavior is similar to that of the n-JFET, except that with $v_{GS} = 0$ the device is off. To make the channel conductive, v_{GS} must be raised above some threshold $V_{GS(on)} > 0$; the greater v_{GS} compared to $V_{GS(on)}$, the more conductive the channel. When operated in a virtual-ground arrangement of the type of Fig. 9.37, an n-MOSFET opens when the gate voltage is low and closes when the gate voltage is high.

If the n-MOSFET is connected in a floating arrangement of the type of Fig. 9.36, the on-state conductivity is no longer uniformly high, but varies with v_I since v_{GS} itself is a function of v_I. The channel is much less conductive during positive than during negative alternations of v_I, and for sufficiently positive values of v_I it may actually turn off. These drawbacks are eliminated by using a pair of complementary MOS (CMOS) FETs, one handling the negative and the other the positive alternations of v_I. The former is an enhancement n-MOSFET, and the latter an enhancement p-MOSFET, whose characteristics are similar to those of the n-MOSFET, except that the turn-on threshold is now negative. So, to make a p-MOSFET conductive, we need $v_{GS} < V_{GS(on)} < 0$; the lower v_{GS} compared to $V_{GS(on)}$, the more conductive the channel. For proper operation, the p-MOSFET must be driven in antiphase with respect to the n-MOSFET. As shown in Fig. 9.39*a* for the case of symmetric power supplies, this drive is provided by an ordinary CMOS inverter.

When C/O is high, the gate of n-MOSFET M_n is high and that of p-MOSFET M_p is low, turning both devices on. As depicted in Fig. 9.39*b*, M_n offers low resistance only over the lower portion of the signal range, and M_p only over the upper portion. However, as a team, they offer a combined parallel resistance that is reasonably low throughout the entire range $V_{SS} \leq v_I \leq V_{DD}$. Finally, when C/O is low, both FETs are off and signal transmission is inhibited.

Also called a *transmission gate,* the basic configuration of Fig. 9.39*a* is available in a variety of versions and performance ratings. Two of the oldest examples are the CD4066 quad bilateral switch and the CD4051 eight-channel multiplexer/demultiplexer, originally introduced by RCA. The 4051 also provides

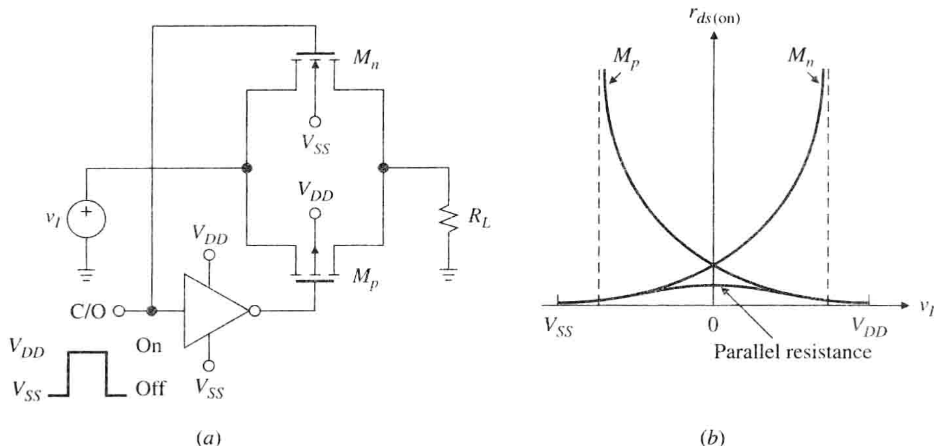


FIGURE 9.39
CMOS transmission gate and its dynamic resistance as a function of v_I.

logic-level translation to allow the switches to work with bipolar analog signals while accepting unipolar logic levels. A wide variety of other MOS-switch products can be found by consulting the data books.

9.6 PEAK DETECTORS

The function of a *peak detector* is to capture the peak value of the input and yield $v_O = v_{I(\text{peak})}$. To achieve this goal, v_O is made to track v_I until the peak value is reached. This value is then held until a new, larger peak comes along, in which case the circuit will update v_O to the new peak value. Figure 9.40*a* shows an example of input and output waveforms. Peak detectors find application in test and measurement instrumentation.

From the above description we identify the following four blocks: (*a*) an analog memory to hold the value of the most recent peak—this is the capacitor, whose ability to store charge makes it act as a voltage memory, as per $V = Q/C$; (*b*) a

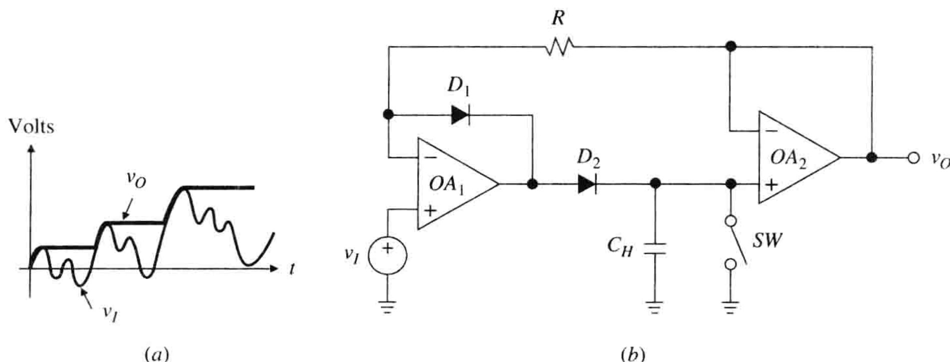


FIGURE 9.40
Peak-detector waveforms and circuit diagram.

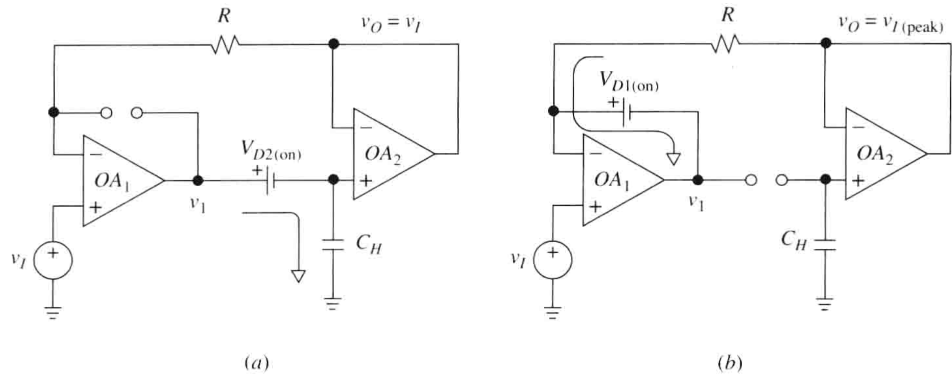


FIGURE 9.41
Peak-detector equivalents during (*a*) the track mode, and (*b*) the hold mode.

unidirectional current switch to further charge the capacitor when a new peak comes along: this is the diode; (*c*) a device to force the capacitance voltage to track the input voltage when a new peak comes along: this is the voltage follower; (*d*) a switch to periodically reinitialize v_O to zero: this is accomplished with a FET discharge switch in parallel with the capacitor.

In the circuit of Fig. 9.40*b* the above tasks are performed, respectively, by C_H, D_2, OA_1, and SW. The function of OA_2 is to buffer the capacitor voltage to prevent discharge by R and by any external load. Moreover, D_1 and R prevent OA_1 from saturating after a peak has been detected, and thus speed up recovery when a new peak comes along. The circuit operates as follows.

With the arrival of a new peak, OA_1 swings its output v_1 positive, turning D_1 off and D_2 on as shown in Fig. 9.41*a*. OA_1 uses the feedback path D_2-OA_2-R to maintain a virtual short between its inputs. Since no current flows through R, the result is that v_O will track v_I. During this mode, aptly called the *track mode,* OA_1 sources current to charge C_H via D_2, and its output rides a diode drop above v_O, or $v_1 = v_O + V_{D2(on)}$.

After peaking, v_I starts to decrease, causing the output of OA_1 also to decrease. Consequently, D_2 goes off and D_1 goes on, thus providing an alternative feedback path for OA_1, as depicted in Fig. 9.41*b*. By the virtual-short concept, the output of OA_1 now rides a diode drop below v_I, or $v_1 = v_I - V_{D1(on)}$. During this mode, called the *hold mode,* the capacitor voltage remains constant, and the function of R is to provide a current path for D_1.

We observe that placing D_2 and OA_2 within the feedback path of OA_1 eliminates any errors due to the voltage drop across D_2 and the input offset voltage of OA_2. All that is required at the input of OA_2 is a sufficently low input bias current to minimize capacitance discharge between peaks. The requirements of OA_1 are a sufficiently low dc input error, and a sufficiently high output-current capability to charge C_H during fast peaks. Moreover, OA_1 may need to be stabilized against the feedback-loop pole introduced by r_{o1} and C_H, and that introduced by OA_2. This is usually achieved by connecting suitable compensation capacitors in parallel with D_1 and R. Typically, R is on the order of a few kiloohms, and the compensation capacitances on the order of a few tens of picofarads.

It is readily seen that reversing the diode directions causes the circuit to detect the negative peaks of v_I.

Voltage Droop and Sagback

During the hold mode, v_O should remain rigorously constant. In practice, because of leakage currents, the capacitor will slowly charge or discharge, depending on leakage polarity. Leakage stems from various sources, namely, from diode, capacitor, and reset-switch leakage; printed-circuit board leakage; and the input bias current of OA_2. Using the capacitance law $i = C\,dv/dt$ and denoting the net capacitance leakage as I_L, we define the *voltage droop rate* as

$$\frac{dv_O}{dt} = \frac{I_L}{C_H} \tag{9.23}$$

For instance, a 1-nA leakage current through a 1-nF capacitance produces a voltage droop rate of $10^{-9}/10^{-9} = 1$ V/s = 1 mV/ms. Droop is minimized by reducing the individual leakage components.

The most crucial limitations of a practical capacitor in analog memory applications are *leakage* and *dielectric absorption.* Leakage causes the device to slowly discharge when in the hold mode; dielectric absorption causes the new voltage to creep back toward the previous voltage after the capacitance is subjected to a rapid voltage change. This sagback effect, stemming from charge storage phenomena in the bulk of the dielectric, can be modeled with a series of internal *R-C* stages, each in parallel with C_H. Referring to the first-order model of Fig. 9.42*a,* we observe that even though C_H discharges almost instantaneously when *SW* is closed, C_{DA} will retain some charge because of the series resistance R_{DA}. After *SW* is opened, C_{DA} will transfer part of its charge back to C_H to achieve equilibrium, thus causing the sagback effect depicted in Fig. 9.42*b*. Though more than one time constant may intervene in the sag, a single time constant is often sufficient to characterize the sag, with C_{DA} typically one or more orders of magnitude smaller than C_H, and a time constant ranging from fractions of a millisecond to fractions of a second. Capacitor types are available with low leakage and low dielectric absorption. These include polystyrene, polypropylene, and Teflon types.[3]

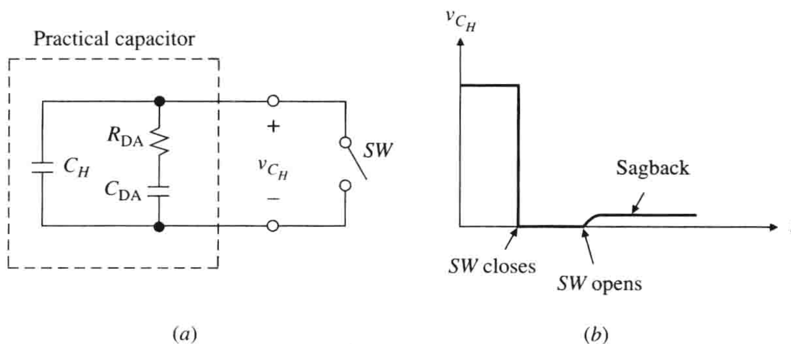


FIGURE 9.42
(*a*) Circuit model for dielectric absorption, and (*b*) the sagback effect.

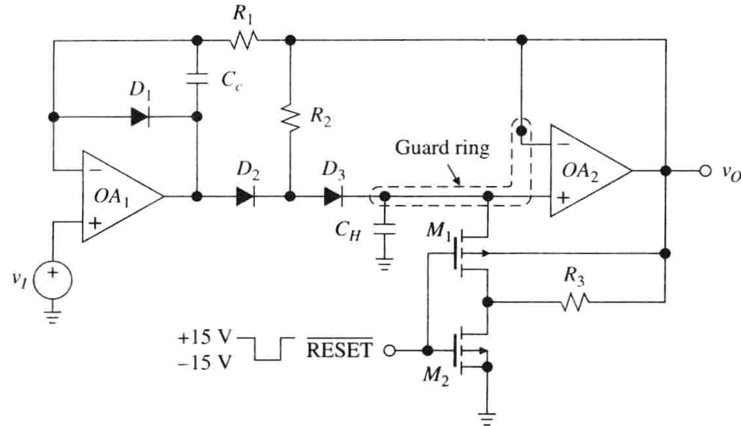


FIGURE 9.43
Peak detector for extended hold.

Printed-circuit board leakage is minimized by input guarding techniques of the type discussed in Section 5.3. In the present circuit the ring is driven by v_O and is made to surround all traces associated with the noninverting input of OA_2, as shown in the practical example of Fig. 9.43.

A FET-input op amp is often chosen for OA_2 to take advantage of its low-input-bias-current characteristics. However, this current doubles with about every 10 °C increase, so if an extended range of operating temperatures is anticipated, a BJT-input op amp with ultralow-input bias current may be preferable.

When reverse biased, a diode draws a leakage current that also doubles with every 10 °C increase. The circuit of Fig. 9.43 eliminates the effect of diode leakage by using a third diode D_3 and the pullup resistance R_2. During the track mode, the D_2-D_3 pair acts as a unidirectional switch, but with a voltage drop twice as large. During the hold mode, R_2 pulls the anode of D_3 to the same potential as the cathode, thus eliminating D_3's leakage; the reverse bias is sustained solely by D_2.

A similar technique can be used to minimize reset-switch leakage. In the example shown, this switch is implemented with two 3N163 enhancement *p*-MOSFETs (Siliconix). Applying a negative pulse to their gates turns both FETs on and also discharges C_H. Upon pulse removal both FETs go off; however, with R_3 pulling the source of M_1 to the same potential as the drain, M_1's leakage is eliminated; the switch voltage is sustained solely by M_2. If TTL compatibility is desired, one can use a suitable voltage-level translator, such as the DH0034.

A good choice for the op amps of Fig. 9.43 is a dual JFET-input device such as the precision, high-speed OP249 op amp. The diodes can be any general-purpose devices, such as the 1N914 or 1N4148 types, and suitable values for the various resistances are in the 10-kΩ range. The purpose of C_c, typically in the range of a few tens of picofarads, is to stabilize the capacitively loaded op amp OA_1 during the track mode. C_H should be large enough to reduce the effect of leakage, yet small enough to allow for its rapid charge during fast peaks. A reasonable compromise is typically in the 1-nF range.

Speed Limitations

Peak-detector speed is limited by the slew rates of its op amps as well as the maximum rate at which OA_1 can charge or discharge C_H. The latter is I_{sc1}/C_H, where I_{sc1} is the short-circuit output current of OA_1. For instance, with $C_H = 0.5$ nF, an op amp having $SR_1 = 30$ V/μs and $I_{sc1} = 20$ mA gives $I_{sc1}/C_H = 40$ V/μs, indicating that speed is limited by SR_1. However, with $C_H = 1$ nF, we get $I_{sc1}/C_H = 20$ V/μs, indicating that speed is now limited by I_{sc1}. The output current drive of OA_1 can be boosted by replacing D_3 with the *B-E* junction of an *npn* BJT, whose collector is returned to V_{CC} via a series resistance on the order of 10^2 Ω to limit current spikes below a proper safety level.

9.7 SAMPLE-AND-HOLD AMPLIFIERS

It is often necessary to capture the value of a signal in response to a suitable logic command, and hold it until the arrival of a new capture command. We have been exposed to this concept in Chapter 5 in connection with autozeroing amplifiers, where the signal in question is an offset-nulling voltage. Other examples will be encountered in Chapter 12 in connection with A-D and D-A converters.

A *sample-and-hold amplifier* (SHA) is a circuit in which the value of the input signal is captured instantaneously, as shown in Fig. 9.44*a*. Though mathematically convenient in sampled-data theory, instantaneous capture is unfeasible because of inherent dynamic limitations of physical circuits. Rather, a practical circuit is made to track the input for a prescribed time interval, and then hold its most recent value for the remainder of the cycle. The timing of the *track-and-hold amplifier* (THA) is shown in Fig. 9.44*b*. In spite of the obvious differences between the diagrams, engineers use the designations SHA and THA interchangeably.

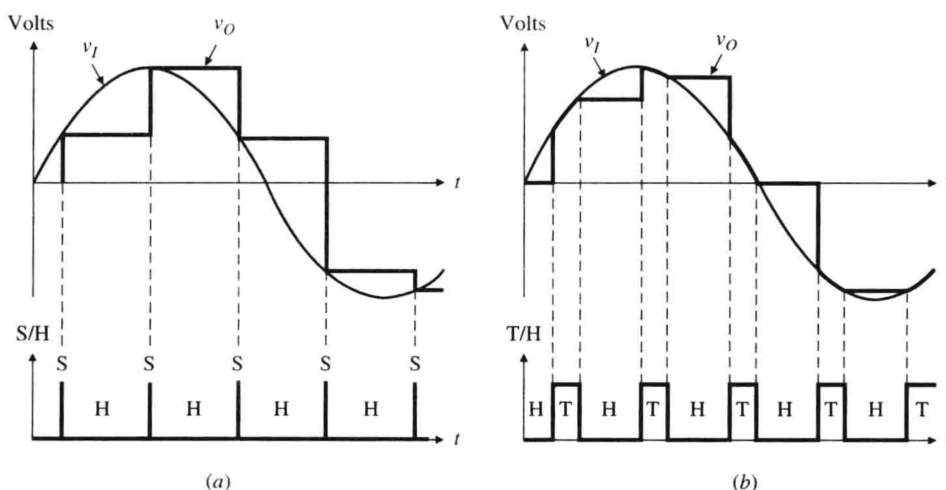


FIGURE 9.44
Idealized responses of (*a*) the sample-and-hold amplifier (SHA), and (*b*) the track-and-hold amplifier (THA).

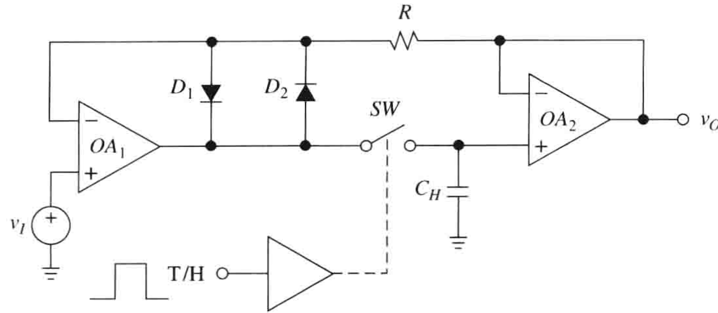


FIGURE 9.45
Basic track-and-hold amplifier.

Figure 9.45 shows one of the most popular THA topologies. The circuit is reminiscent of the peak detector, except for the replacement of the diode switch with an externally controlled bidirectional switch to charge as well as discharge C_H, depending on the case. The circuit operates as follows.

During the *track* mode, SW is closed to create the feedback path SW-OA_2-R around OA_1. Due to the low voltage drop across SW, both diodes are off, indicating a 0-V drop across R. OA_1 thus acts as a voltage follower, providing C_H with whatever current it takes to make v_O track v_I.

During the *hold* mode, SW is opened, allowing C_H to retain whatever voltage it had at the instant of switch aperture; OA_2 then buffers this voltage to the outside. The function of D_1 and D_2 is to prevent OA_1 from saturating, and thus facilitate OA_1's recovery when a new track command is received.

The switch is usually a JFET, a MOSFET, or a Schottky diode bridge, and is equipped with a suitable driver to make the T/H command TTL- or CMOS-compatible. The main requirements of OA_1 are (*a*) low-input dc error, (*b*) adequate output current capability to rapidly charge or discharge C_H, (*c*) high open-loop gain to minimize the gain error and errors due to the voltage drop across SW and OA_2's input offset voltage, and (*d*) proper frequency compensation for sufficiently fast dynamics and settling characteristics. Compensation is often implemented with a bypass capacitance of a few tens of picofarads in parallel with the diodes. The requirements of OA_2 are (*a*) low-input bias current to minimize droop, and (*b*) adequately fast dynamics. As in the peak-detector case, C_H should be a low-leakage, low-dielectric-absorption capacitor, such as Teflon or polystyrene.[3] Its value is chosen as a compromise between low droop and rapid charge/discharge times.

The basic THA of Fig. 9.45 can be implemented with individual op amps and passive components, or it can be purchased as a self-contained monolithic IC. A popular example is the LM398 BiFET THA.

THA Performance Parameters

In the track mode, a THA is designed to behave like an ordinary amplifier, so its performance is characterized in terms of the dc and gain errors, the dynamics, and

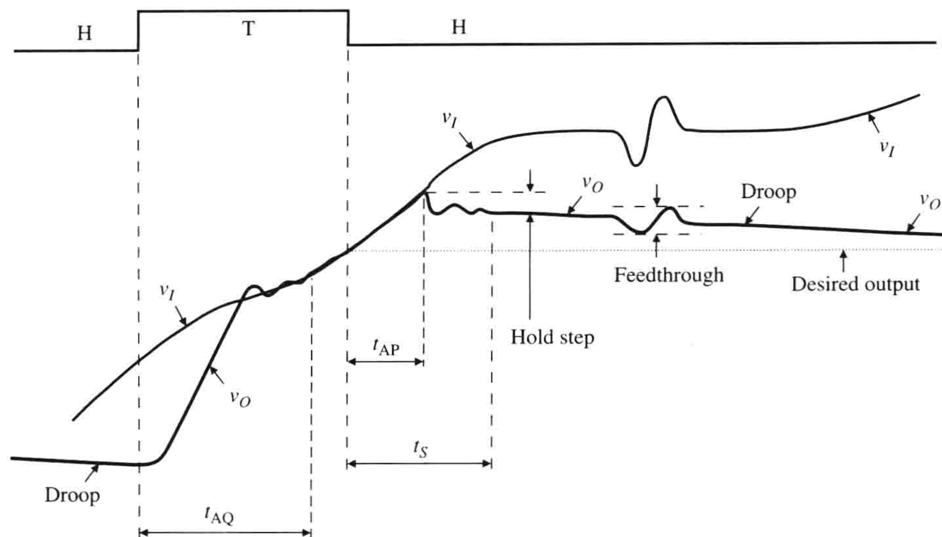


FIGURE 9.46
THA terminology.

other parameters peculiar to amplifiers. However, during the transition from the track to the hold mode and vice versa, as well as during the hold mode itself, performance is characterized by specifications peculiar to THAs. The following list uses the expanded timing diagram of Fig. 9.46 as a guideline.

1. *Acquisition Time* (t_{AQ}). Following the track command, v_O starts slewing toward v_I, and t_{AQ} is the time it takes for v_O to begin tracking v_I within a specified error band after the inception of the track command. This includes propagation delays through the switch driver and the switch, and delays due to slew-rate limiting and settling times of the op amps. The acquisition time increases with the step magnitude as well as the narrowness of the error band. Usually t_{AQ} is specified for a 10-V step and for error bands of 1%, 0.1%, and 0.01% of full scale. The input must be fully acquired before switching to the hold mode.
2. *Aperture Time* (t_{AP}). Because of propagation delays through the driver and the switch, v_O will cease tracking v_I some time after the inception of the hold command. This is the aperture time. The hold command would have to be advanced by t_{AP} for precise timing.
3. *Aperture Uncertainty* (Δt_{AP}). Also called *aperture jitter,* it represents the variation in aperture time from sample to sample. If t_{AP} is compensated for by advancing the hold command by t_{AP}, then Δt_{AP} establishes the ultimate timing error and, hence, the maximum sampling frequency for a given resolution. Aperture jitter results in an output error $\Delta v_O = (dv_I/dt)\Delta t_{AP}$, indicating that the actual sampled waveform can be viewed as the sum of an ideally sampled waveform and a noise component. The signal-to-noise ratio of an otherwise ideal sampling circuit with a sinusoidal input of frequency f_i is[4]

$$\text{SNR} = -20\log_{10}[2\pi f_i \Delta t_{AP(\text{rms})}] \tag{9.24}$$

where $\Delta t_{AP(rms)}$ is the rms value of Δt_{AP}, and the latter is assumed uncorrelated with v_I. Typically, Δt_{AP} is an order of magnitude smaller than t_{AP}, and t_{AP}, in turn, is one to two orders of magnitude smaller than t_{AQ}.

4. *Hold Mode Settling Time (t_S).* After the inception of the hold command, it takes some time to settle within a specified error band, such as 1%, 0.1%, or 0.01%. This is the hold mode settling time.
5. *Hold Step.* Because of parasitic switch capacitances, when the circuit is switched to the hold mode there is an unwanted charge transfer between the switch driver and C_H, causing a change in the voltage across C_H. The corresponding change Δv_O is referred to as *hold step, pedestal error,* or *sample-to-hold offset.*
6. *Feedthrough.* When in the hold mode, v_O should be independent of any variations in v_I. In practice, because of stray capacitance across SW, there is a small amount of ac coupling from v_I to v_O called *feedthrough.* This capacitance forms an ac voltage divider with C_H, so an input change Δv_I causes an output change $\Delta v_O = [C_{SW}/(C_{SW} + C_H)]\Delta v_I$, where C_{SW} is the capacitance across the switch. The *feedthrough rejection ratio*

$$\text{FRR} = 20 \log \frac{\Delta v_I}{\Delta v_O} \tag{9.25}$$

gives an indication of the amount of stray coupling. For example, if FRR = 80 dB, a hold mode change $\Delta v_I = 10$ V results in $\Delta v_O = \Delta v_I/10^{80/20} = 10/10^4 =$ 1 mV.

7. *Voltage Droop.* THAs are subject to the same droop limitations as peak detectors. Droop is of special concern when C_H must be kept low to ensure a fast acquisition.

In the case of a JFET switch, feedthrough is due to the drain-source capacitance C_{ds}, and the hold step is due to the gate-drain capacitance C_{gd}. (For discrete devices, these capacitances are typically in the picofarad range.) As the driver pulls the gate from near v_O to near V_{EE}, it removes the charge $\Delta Q \cong C_{gd}(V_{EE} - v_O)$ from C_H, causing a hold step

$$\Delta v_O \cong \frac{C_{gd}}{C_{gd} + C_H}(V_{EE} - v_O) \tag{9.26}$$

This step varies with v_O. For example, with $C_H = 1$ nF and $V_{EE} = -15$ V, the hold step for every picofarad of C_{gd} is about -15 mV/pF for $v_O = 0$, -20 mV/pF for $v_O = 5$ V, and -10 mV/pF for $v_O = -5$ V. A C_{gd} of just a few picofarads can cause intolerable errors!

There are various techniques for minimizing the signal-dependent hold step. One such technique is to implement the switch with the CMOS transmission gate of Fig. 9.39*a*. Since the two FETs are driven in antiphase, one FET will inject and the other will remove charge, and if their geometries are properly scaled, the two charges will cancel each other out.

An alternative technique[5] is depicted in Fig. 9.47. As the circuit goes into hold, OA_4 produces a positive-going output swing that, by the superposition principle, depends on v_O as well as on the negative step on the switch gate. This swing is

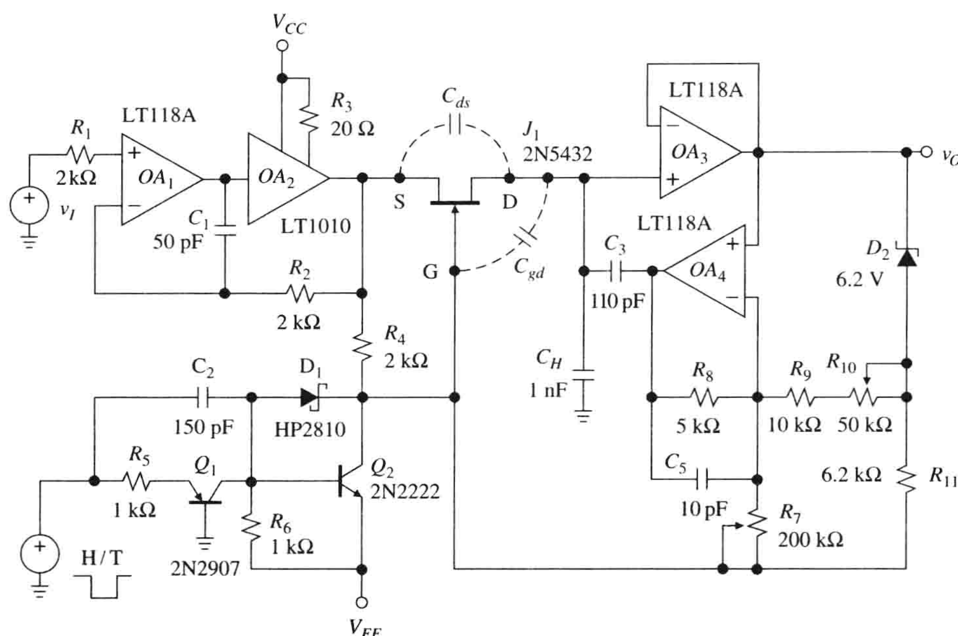


FIGURE 9.47
A 5-MHz THA with charge-transfer compensation to minimize the hold step. (Courtesy of Linear Technology.)

designed to inject into C_H, via C_3, a charge packet of magnitude equal to that removed via C_{gd}, thus resulting in a net charge transfer of zero. The hold step is made independent of v_O with R_7, and adjusted to zero via R_{10}. To calibrate the circuit, adjust R_7 for equal hold steps with $v_I = \pm 5$ V; then, null the residual offset via R_{10}.

To achieve high speed, the circuit utilizes fast op amps and boosts OA_1 with the LT1010 fast power buffer to rapidly charge and discharge C_H in the track mode. Moreover, by using local feedback around the OA_1-OA_2 pair, the settling dynamics of the input and output stages are kept separate and simpler. With OA_3 no longer inside the control loop, its input offset voltage is no longer irrelevant; however, its offset as well as that of the input buffer are automatically compensated for during calibration of R_{10}. For long hold periods, OA_3 can be replaced by a FET-input device such as the LF356 to reduce droop.

Charge compensation can be simplified considerably if the switch is operated in a virtual-ground arrangement. This is the case with THAs of the *integrating type*,[4,6] so-called because the holding capacitor is placed in the feedback path of the output amplifier, as exemplified in Fig. 9.48. Since the switch always sees a virtual ground, the charge removed from the summing junction via C_{gd} is constant regardless of v_O. Consequently, the hold step appears as a constant offset that can easily be nulled by standard techniques, such as adjusting the offset of OA_1, as shown. With an easier-to-compensate hold step, the holding capacitance can be reduced significantly to achieve faster acquisition times. The HA-5330 high-speed monolithic THA uses a 90-pF holding capacitance to achieve $t_{AQ} = 400$ ns to 0.01%.

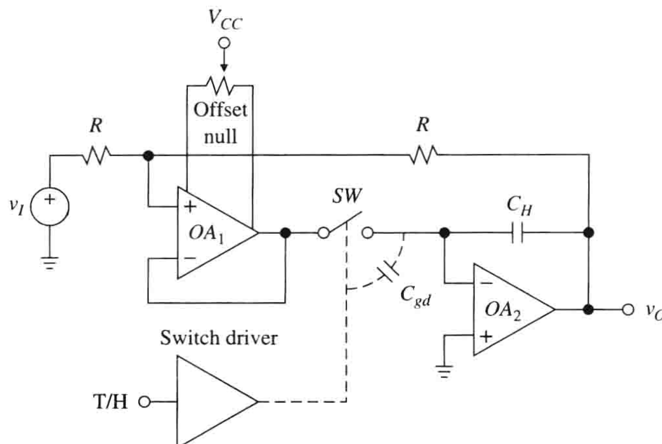


FIGURE 9.48
Integrating-type THA.

Figure 9.49 shows an improved integrating-type THA that simultaneously optimizes droop, hold step, and feedthough. During the track mode, SW_3 is open while SW_1 and SW_2 are closed. In this mode the circuit operates just as in Fig. 9.48, with v_O slewing toward $-v_I$. During the hold mode, SW_1 and SW_2 are open while SW_3 is closed, causing the circuit to hold whatever voltage it acquired during sampling. Note, however, that by grounding the input to the buffer via SW_3, any variations in v_I are muzzled, thus improving the FRR significantly. Moreover, since both SW_1 and SW_2 experience a voltage drop very close to zero, switch leakage is virtually eliminated. The main source of leakage is now the input bias current of *OA*. However, returning its noninverting input to a dummy capacitance C of size equal to C_H produces a hold step and a droop that, to a first approximation, will cancel out the hold step and droop of C_H. An example of a THA utilizing this technique is the SHC803/804, whose typical ambient-temperature ratings are: $t_{AQ} = 250$ ns and $t_S = 100$ ns, both to 0.01%; $t_{AP} = 15$ ns; $\Delta t_{AP} = \pm 15$ ps; FRR $= \pm 0.005\%$, or 86 dB; hold mode offset $= \pm 2$ mV; droop rate $= \pm 0.5\ \mu\text{V}/\mu\text{s}$.

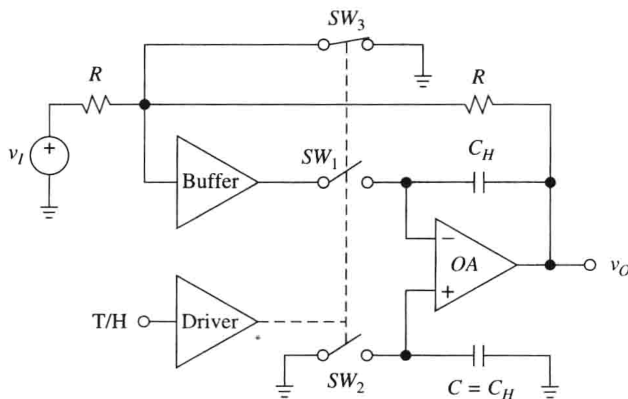


FIGURE 9.49
Improved THA. (Switch settings shown for the hold mode.)

THAs are available from a variety of sources and in a wide range of performance specifications and prices. Consult the catalogs to familiarize yourself with the available products.

PROBLEMS

9.1 Voltage comparators

9.1 (*a*) Using a 311 comparator powered from ±5-V regulated supplies, design a threshold detector that gives $v_O \cong 0$ for $v_I > 1$ V, and $v_O \cong +5$ V for $v_I < 1$ V. (*b*) Repeat, but with $v_O \cong +5$ V for $v_I > -1$ V, and $v_O \cong -5$ V for $v_I < -1$ V. (*c*) Repeat, but with $v_O \cong 2.5$ V for $v_I > 0$, and $v_O \cong -2.5$ V for $v_I < 0$.

9.2 Comparator applications

9.2 The thermal characteristic of a certain class of thermistors can be expressed as $R(T) = R(T_0)\exp[B(1/T - 1/T_0)]$, where T is absolute temperature, T_0 is some reference temperature, and B is a suitable constant, all three parameters being in kelvins. Using a single comparator of the 339 type and a thermistor having $R(25\ ^\circ\text{C}) = 100$ kΩ and $B = 4000$ K, design a bridge comparator circuit that gives $v_O = V_{OH}$ for $T > 100\ ^\circ$C, and $v_O = V_{OL}$ for $T < 100\ ^\circ$C. Assuming 10% component tolerances, make provision for the exact adjustment of the setpoint, and outline the calibration procedure.

9.3 Using an op amp, two comparators of the 339 type, a 2N2222 *npn* BJT, and resistors as needed, design a circuit that accepts a data input v_I and a control input V_T, $0 < V_T \le 2.5$ V, and causes a 10-mA LED to glow whenever $-V_T \le v_I \le V_T$. Assume ±5-V power supplies.

9.4 Using two comparators of the 339 type and a thermistor of the type of Problem 9.2 with $R(25\ ^\circ\text{C}) = 10$ kΩ and $B = 4000$ K, design a circuit that yields $v_O \cong 5$ V for $0\ ^\circ\text{C} \le T \le 5\ ^\circ$C, and $v_O \cong 0$ V otherwise. Assume a single 5-V regulated supply.

9.5 Show that the window detector of Fig. P9.5 has a window whose center is controlled by v_1 and whose width is controlled by v_2; then sketch and label the VTC if $v_1 = 3$ V and $v_2 = 1$ V.

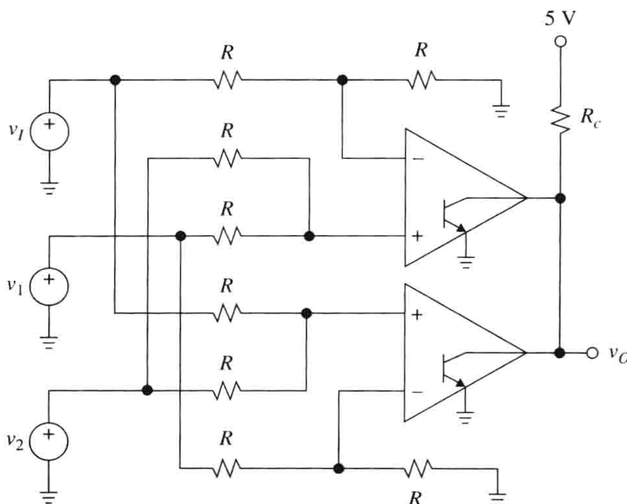


FIGURE P9.5

9.6 Using three comparators of the 339 type, an LM385 2.5-V reference diode, an HLMP-4700 LED of the type of Example 9.2, and resistors as needed, design a circuit that monitors a 15-V $\pm$ 5% power supply and causes the LED to glow whenever the supply is within range.

9.7 Using an LM385 2.5-V reference diode, an LM339 quad comparator, and four HLMP-4700 LEDs of the type of Example 9.2, design a 0-V to 4-V bar graph meter. The circuit must have an input impedance of at least 100 kΩ and must be powered from a single 5-V supply.

9.8 Using a 311 comparator powered from $\pm$15-V regulated supplies, design a circuit that accepts a triangular wave with peak values of $\pm$10 V, and generates a square wave with peak values of $\pm$5 V and duty cycle D variable from 5% to 95% by means of a 10-kΩ potentiometer.

9.3 Schmitt triggers

9.9 In the circuit of Fig. 9.20a let v_I be a triangular wave of $\pm$10 V peak values and let $\pm V_{sat} = \pm 13$ V. Modify the circuit so that the phase of its square-wave output, relative to that of the input, is variable from 0° to 90° by means of a 10-kΩ potentiometer. Show the input and output waveforms when the wiper is in the middle.

9.10 (a) Derive Eq. (9.14). (b) Specify suitable resistances in the circuit of Fig. 9.23 to achieve $V_{OL} = 0$ V, $V_{OH} = 5$ V, $V_{TL} = 1.5$ V, and $V_{TH} = 2.5$ V with $V_{CC} = 5$ V. Try minimizing the effect of the input bias current.

9.11 Assuming $V_{D(\text{on})} = 0.7$ V and $\pm V_{sat} = \pm 13$ V, sketch and label the VTC of the inverting Schmitt trigger of Fig. P9.11.

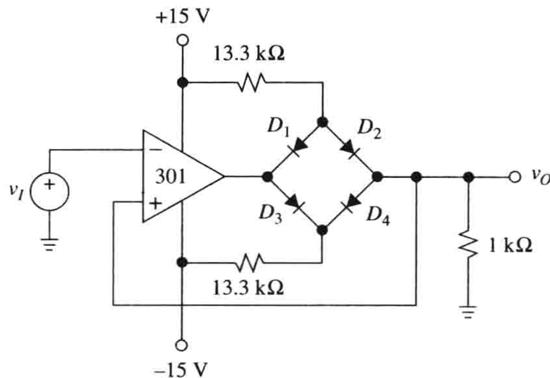


FIGURE P9.11

9.12 (a) Assuming the op amp of Fig. 9.20a saturates at $\pm$13 V, sketch and label the VTC if a resistance $R_3 = 33$ kΩ is connected between the nodes labeled v_P and -15 V. (b) Suitably modify the circuit of Fig. 9.21a so that it gives $V_{TL} = 1$ V and $V_{TH} = 2$ V.

9.13 (a) Using CMOS inverters of the type shown in Fig. 10.11, along with resistances in the 10-kΩ to 100-kΩ range, design a noninverting Schmitt trigger with $V_{TL} = (1/3)V_{DD}$ and $V_{TH} = (2/3)V_{DD}$; assume $V_T = 0.5V_{DD}$. (b) Modify the circuit so that $V_{TL} = (1/5)V_{DD}$

and $V_{TH} = (1/2)V_{DD}$. (*c*) How would you turn the preceding circuits into Schmitt triggers of the inverting type?

9.14 Suitably modify the circuit of Problem 9.2 to ensure a hysteresis of ±0.5 °C. Outline the calibration procedure.

9.15 In the Schmitt trigger of Fig. 9.20*a* let the input v_I be applied to the inverting-input pin via a voltage divider made up of two 10-kΩ resistances, let R_1 be replaced by the series combination of two 4.3-V Zener diodes connected back to back anode with anode, and let the output v_O be obtained at the node where R_2 joins the Zener network. Draw the circuit. Hence, assuming a 0.7-V forward-bias diode voltage drop, sketch and label the VTC.

9.16 In the circuit of Fig. P1.18 let the source be a variable source, denoted as i_I, and let the op amp saturate at ±10 V. (*a*) Sketch and label v_O versus i_I for i_I variable over the range -1 mA $\leq i_I \leq 1$ mA. (*b*) Repeat, but with a 2-kΩ resistor in parallel with i_I, and for i_I variable over the range -2 mA $\leq i_I \leq 2$ mA.

9.17 A circuit consists of a 311 comparator and three equal resistors, $R_1 = R_2 = R_3 =$ 10 kΩ. The 311 is powered between 15 V and ground, and has $V_{EE(\text{logic})} = 0$. Moreover, R_1 is connected between the 15-V supply and the noninverting-input pin, R_2 between the noninverting-input pin and the open-collector output pin, and R_3 between the open-collector output pin and ground. Moreover, the input v_I is applied to the comparator's inverting-input pin. Draw the circuit, and sketch and label its VTC if the output v_O is obtained from: (*a*) the node where R_1 joins R_2; (*b*) the node where R_2 joins R_3.

9.18 Consider the circuit obtained by removing R, C, and *OA* from Fig. 10.19*a*. What is left then is a noninverting Schmitt trigger, whose input is the node labeled as v_{TR}, and whose output is the node labeled as v_{SQ}. Sketch and label its VTC if $R_1 = 10$ kΩ, $R_2 = 13$ kΩ, $R_3 = 4.7$ kΩ, and the Zener diode is a 5.1-V device; assume forward-bias diode voltage drops of 0.7 V.

9.19 (*a*) Using a 339 comparator powered from a 5-V regulated supply, design an inverting Schmitt trigger such that $V_{OL} = 2.0$ V, $V_{OH} = 3.0$ V, $V_{TL} = 2.0$ V, and $V_{TH} = 3.0$ V. (*b*) Modify your circuit so that $V_{TL} = 0$ V and $V_{TH} = 5.0$ V.

9.4 Precision rectifiers

9.20 Sketch and label the VTC of the circuit of Fig. 9.29*a* if $R_2 = 2R_1$ and the noninverting input of the op amp is lifted off ground and returned to a -5-V reference voltage. Next, sketch and label v_O if v_I is a triangular wave with peak values of ±10 V.

9.21 Sketch and label the VTC of the circuit of Fig. 9.29*a* if $R_1 = R_2 = 10$ kΩ, and a third resistance $R_3 = 150$ kΩ is connected between the +15-V supply and the inverting-input pin of the op amp. (*b*) Repeat, but with the diode polarities reversed.

9.22 One side of a 10-kΩ resistance is driven by a source v_I, and the other side is left floating. Denoting the voltage at the floating side as v_O, use a superdiode circuit to implement a *variable precision clamp,* that is, a circuit that gives $v_O = v_I$ for $v_I \leq V_{\text{clamp}}$ and $v_O = V_{\text{clamp}}$ for $v_I \geq V_{\text{clamp}}$, where V_{clamp} is a continuously adjustable voltage from 0 to 10 V by means of a 100-kΩ pot. Assume ±15-V regulated supplies. List advantages and drawbacks of your circuit.

9.23 Suitably modify the FWR of Fig. 9.30 so that, when fed with a triangular wave of ±5-V peak values, it gives a triangular wave of ±5-V peak values, but twice the frequency. Assume ±10-V regulated power supplies.

9.24 (*a*) Use PSpice to investigate the behavior of the FWR of Fig. 9.30 in response to a sinusoidal input of ±5-V peak values. Use LM324 op amps powered from ±9-V supplies, 1N4148 diodes, and $R = 10\ \text{k}\Omega$ throughout, along with $A = 1$. Display v_I, v_O, and v_{HW}, starting out with an input frequency of 1 kHz. Rerun PSpice for increasing values of the input frequency until the circuit begins to distort. Based on what you see, can you explain the reasons?

9.25 Assuming $R_1 = R_2 = R_4 = 10\ \text{k}\Omega$ and $R_3 = 20\ \text{k}\Omega$ in the FWR of Fig. 9.31, find all node voltages for $v_I = 10$ mV, 1 V, and −1 V. For a forward-biased diode, assume $v_D = (26\ \text{mV}) \ln[i_D/(20\ \text{fA})]$.

9.26 Discuss the effect of resistance mismatches in the FWR of Fig. 9.31, and derive an expression for $100|(A_p - A_n)/A|$. Compare with the FWR of Fig. 9.30, and comment.

9.27 Consider the circuit obtained from that of Fig. 9.31 by grounding the left terminals of R_1 and R_4, lifting the noninverting input of OA_1 off ground and driving it with source v_I. (*a*) Show that the modified circuit gives $v_O = A_p v_I$ for $v_I > 0$ and $v_O = -A_n v_I$ for $v_I < 0$, where $A_p = 1 + (R_2 + R_3)/R_1$ and $A_n = R_3/R_2$. (*b*) Specify component values for $v_O = 5|v_I|$. List advantages and disadvantages of this circuit.

9.28 Consider the circuit obtained from that of Fig. 9.31 by removing R_1, grounding the left terminal of R_4, lifting the noninverting input of OA_1 off ground and driving it with source v_I. Analyze the modified circuit if $R_2 = R_3 = R$. Afterward, discuss the implications of mismatched resistances.

9.29 (*a*) Find the VTC of the circuit of Fig. P9.29. (*b*) Assuming $\pm V_{sat} = \pm 13$ V and $V_{D(on)} = 0.7$ V, show all node voltages for $v_I = +3$ V and $v_I = -5$ V. (*c*) List advantages and disadvantages of this circuit.

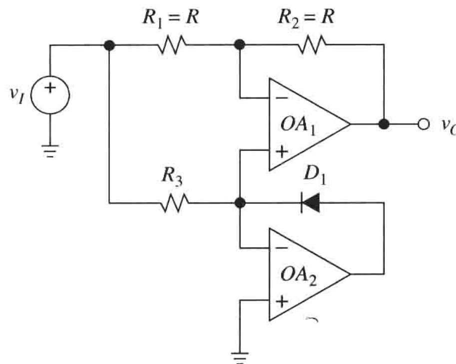


FIGURE P9.29

9.30 The circuit of Fig. 9.30 can be turned into a high-input-impedance FWR by lifting both noninverting inputs off ground, tying them together, and driving them with a common input v_I; moreover, R_4 is removed, and the left terminal of R_1 is grounded. (*a*) Assuming $R_1 = R_2 = R_3 = R$ and $R_5 = 2R$, find the VTC of the modified circuit. (*b*) Assuming $V_{D(on)} = 0.7$ V, show all node voltages for $v_I = +2$ V and $v_I = -3$ V. (*c*) Investigate the effect of mismatched resistances.

9.31 (*a*) Find the VTC of the circuit of Fig. P9.31; then, assuming $V_{D(\text{on})} = 0.7$ V, show all node voltages for $v_I = +1$ V and $v_I = -3$ V. (*b*) Suitably modify the circuit so that it accepts two inputs v_1 and v_2, and gives $v_O = |v_1 + v_2|$.

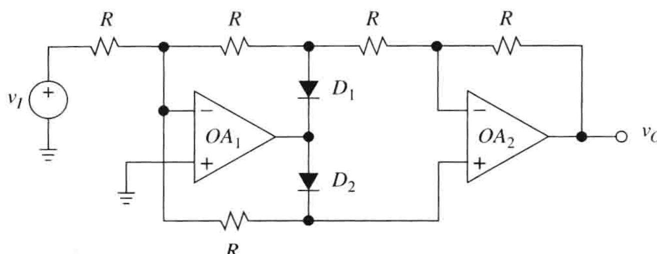


FIGURE P9.31

9.32 Repeat Problem 9.24, but for the circuit of Fig. P9.31.

9.33 Investigate the effect of the input offset voltages V_{OS1} and V_{OS2} of OA_1 and OA_2 in the FWR of Fig. 9.30.

9.5 Analog switches

9.34 Using a 311 comparator, a 2N4391 *n*-JFET, and a 741 op amp, design a circuit that accepts an analog signal v_I and two control signals v_1 and v_2, and yields a signal v_O such that $v_O = 10v_I$ for $v_1 > v_2$, and $v_O = -10v_I$ for $v_1 < v_2$. Assume ±15-V supplies.

9.35 For small values of $|v_{DS}|$, the channel resistance of a MOSFET can be found as $1/r_{ds(\text{on})} \cong k(|v_{GS}| - |V_{GS(\text{on})}|)$, where k is called the *device transconductance parameter,* in amperes per square volt. Assuming ±5-V supplies in the transmission gate of Fig. 9.39*a*, and truly complementary FETS with $k = 100\ \mu\text{A/V}^2$ and $|V_{GS(\text{on})}| = 2.5$ V, find the net switch resistance for $v_I = \pm 5$ V, ±2.5 V, and 0 V. What are the corresponding values of v_O if $R_L = 100\ \text{k}\Omega$?

9.6 Peak detectors

9.36 Consider the circuit obtained from that of Fig. 9.40*b* by returning the noninverting input of OA_1 to ground, and applying the source v_I to the inverting input of OA_1 via a series resistance having the same value as the feedback resistance *R*. Discuss how the modified circuit operates, and show its response to a sinusoidal input of increasing amplitude.

9.37 Design a peak-to-peak detector, that is, a circuit that gives $v_O = v_{I(\max)} - v_{I(\min)}$.

9.38 Using the circuit of Fig. 9.29*a* as a starting point, design a circuit to provide the *magnitude peak-detector* function, $v_O = |v_I|_{\max}$.

9.39 Three superdiodes of the type of Fig. 9.27 are driven by three separate sources v_1, v_2, and v_3, and their outputs are tied together and returned to −15 V via a 10-kΩ resistor. What function does the circuit provide? What happens if the diode polarities are reversed? If the node common to the outputs is returned to the node common to the inverting inputs via a voltage divider?

9.7 Sample-and-hold amplifiers

9.40 Suitably modify the THA of Fig. 9.45 for a gain of 2 V/V. What is the main disadvantage of the modified circuit, and how would you take care of it?

9.41 In the THA of Fig. 9.48 let $C_{gd} = 1$ pF, $C_H = 1$ nF, and let the net leakage current through C_H be 1 nA, flowing from right to left. Assuming $v_I = 1.000$ V, find v_O (*a*) shortly after the circuit is switched to the hold mode, and (*b*) 50 ms later.

9.42 The THA of Fig. P9.42 uses a feedback capacitor $C_F = C_H$ to provide a first-order compensation for the droop due to leakage in C_H. (*a*) Explain how the circuit works. What are the functions of the *p*-channel JFET J_1 and the *n*-channel JFETs J_2 and J_3? (*b*) Assuming an average leakage of 1 nA in each capacitor and a leakage mismatch of 5%, estimate the voltage droop for the case $C_F = C_H = 1$ nF. What would be the leakage if C_F were absent and replaced with a wire?

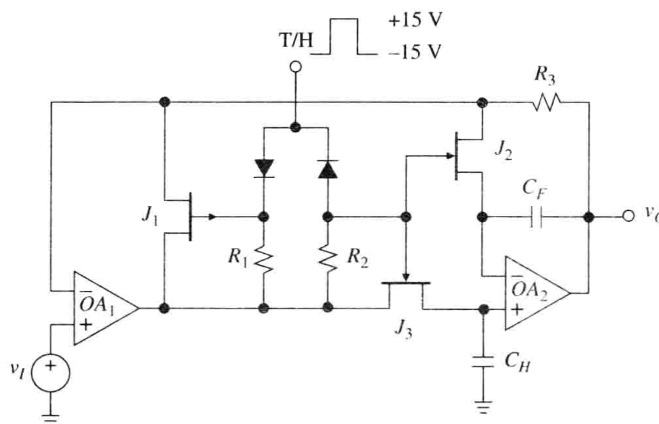


FIGURE P9.42

REFERENCES

1. J. Sylvan, "High-Speed Comparators Provide Many Useful Circuit Functions When Used Correctly," *Analog Dialogue,* Vol. 23, No. 4, Analog Devices, Norwood, MA, 1989.
2. J. Williams, "High-Speed Comparator Techniques," Application Note AN-13, *Linear Applications Handbook Volume I,* Linear Technology, Milpitas, CA, 1990.
3. S. Guinta, "Capacitance and Capacitors," *Analog Dialogue,* Vol. 30, No. 2, Analog Devices, Norwood, MA, 1996.
4. B. Razavi, *Principles of Data Conversion System Design,* IEEE Press, Piscataway, NJ, 1995.
5. R. J. Widlar, "Unique IC Buffer Enhances Op Amp Designs, Tames Fast Amplifiers," Application Note AN-16, *Linear Applications Handbook Volume I,* Linear Technology, Milpitas, CA, 1990.
6. D. A. Johns and K. W. Martin, *Analog Integrated Circuit Design,* John Wiley & Sons, New York, 1997.

10

SIGNAL GENERATORS

到目前为止所研究的电路可以属于处理电路，因为它们都是对已有信号进行处理的。现在要研究这样一类电路，它们本身就是用来产生信号的。尽管有时信号是从传感器中得到的，但是大多数情况下都需要将它们在系统中进行综合。最常见的例子有：用于计时和控制的时钟脉冲；用于信息传输和存储的信号载波；用于显示信息的扫描信号；用于自动检测和测量的测试信号；用于电子音乐和语音合成的音频信号。

信号发生器的功能是产生具有指定特征的（例如频率、幅度、形状以及占空比）波形。有时会通过适当的控制信号将这些特征设计成可在外部编程的，压控振荡器就是一个最典型的例子。通常来说，信号发生器利用某些反馈形式以及像电容那样其特性与时间相关的器件来一起实现。最主要的两类信号发生器就是将要研究的正弦振荡器和弛豫振荡器。

正弦振荡器

这种振荡器利用了系统理论的概念，在复平面的虚轴上创造出一对共轭极点，从而得到持续的正弦振荡。我们曾在第 8 章中十分关注的不稳定性问题现在用来获取期望中的振荡。

一个周期波的正弦纯度是通过它的总谐波畸变来表示的：

$$\text{THD}\ (\%) = 100\sqrt{D_2^2 + D_3^2 + D_4^2 + \cdots} \tag{10.1}$$

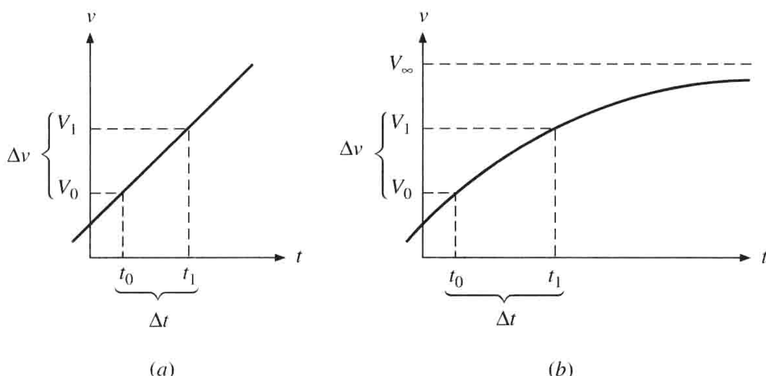
式中，D_k (k = 2，3，4，…) 是给定波形的傅里叶级数中第 k 次谐波与基波的幅值之比。例如，三角波中 $D_k = 1/k^2$，k = 3，5，7，…)，它的 THD=100 × $\sqrt{1/3^4+1/5^4+1/7^4+\cdots} \cong 12\%$，这表明如果把三角波看做正弦波的一个粗略近似，那么它的 THD 就有 12%。另一方面，除基波外，一个纯正弦波形的各次谐波均为零，因此在这种情况下 THD=0%。很显然，设计一个正弦波发生器的目标就是要使 THD 值尽可能低。

弛豫振荡器

这种振荡器使用双稳态器件，如转换器、施密特触发器、逻辑门和触发器等来对电容重复地进行充电与放电。用这种方法得到的典型波形有三角波、锯齿波、指数波形、方波和脉冲波形。在研究过程中，经常需要求出给一个电容充放电达到一个给定的 Δv 所需的时间 Δt。最常见的两种充 / 放电形式是线性和指数形式的。

当以一个恒定的电流 I 驱动电容 C，那么它将会以一个恒定的速率充电或放电，产生一个如图 10.1a 的线性暂态或斜坡。工程师经常用一个易于记忆的关系来描述这个斜坡：

$$C\Delta v = I\Delta t$$



(a) (b)

图 10.1　线性和指数波形

或者"$C\Delta v$ 等于 $I\Delta v$"。这样就可以将恒定速率变化 Δv 所需的时间估算为

$$\Delta t = \frac{C}{I}\Delta v \tag{10.2}$$

当电容 C 通过一个串联电阻 R 进行充电或放电时，会产生一个指数形式的暂态响应。对于图 10.1b，瞬时电容电压为

$$v(t) = V_\infty + (V_0 - V_\infty)\exp[-(t - t_0)/\tau]$$

式中，V_0 是电压初值，V_∞ 是当极限 $t \to \infty$ 所能达到的稳态电压值，而 $\tau = RC$ 是控制暂态响应的时间常数。这个等式的成立与 V_0 和 V_∞ 的值以及极性无关。暂态响应值在 t_1 时刻达到一个指定的中间值 $V_1 = V_\infty+(V_0-V_\infty)\exp[-(t_1-t_0)/\tau]$。将等式两边取自然对数并解出 $\Delta t = t_1-t_0$，就可以将给电容 C 从 V_0 变化 V_1 为所需的充放电时间估算为

$$\Delta t = \tau \ln \frac{V_\infty - V_0}{V_\infty - V_1} \tag{10.3}$$

以后的章节将会经常用到这一等式。

本章要点

本章先研究正弦波振荡器，从文氏电桥振荡器到正交振荡器，最后到其他能在虚轴上自动产生一堆复共轭极点的电路，其中包括第 3 章中讲述的一些滤波器电路。对于不同极点配置的振荡器电路，通过 PSpice 来仿真它们的特性行为。

接着通过一些使用 IC 的典型应用，如 555 计时器，来研究不稳定和双稳态的多谐振荡器。

然后研究三角波和锯齿波发生器，其中也包含电压控制类型的发生器，同时也会提到三角波 – 正弦波转换器。这一部分使用了如 8038 波形发生器、射极耦合的 VCO 和 XR2206 函数产生器等一些常见的 IC。

本章还包括电压 – 频率、频率 – 电压转换器和由它们组成的常规应用。

10.1 SINE WAVE GENERATORS

The sine wave is certainly one of the most fundamental waveforms—both in a mathematical sense, since any other waveform can be expressed as a Fourier combination of basic sine waves, and in a practical sense, since it finds extensive use as a test, reference, and carrier signal. In spite of its simplicity, its generation can be a challenging task if near-purity is sought. The op amp circuits that have gained the most prominence in sine wave generation are the *Wien-bridge oscillator* and the *quadrature*

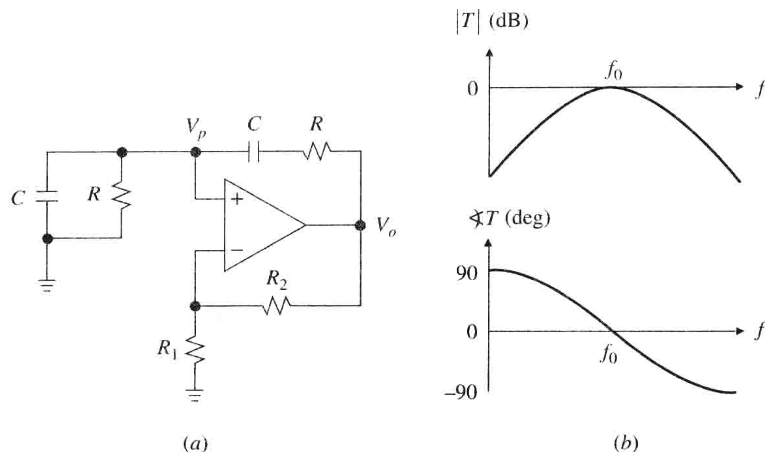


FIGURE 10.2
Wien-bridge circuit and its loop gain $T(jf)$ for the case $R_2/R_1 = 2$.

oscillator, to be discussed next. Another technique, based on the conversion of the triangular to the sine wave, will be discussed in Section 10.4.

Basic Wien-Bridge Oscillator

The circuit of Fig. 10.2*a* uses both negative feedback, via R_2 and R_1, and positive feedback, via the series and parallel *RC* networks. Circuit behavior is strongly affected by whether positive or negative feedback prevails. The components of the *RC* networks need not be equal-valued; however, making them so simplifies analysis as well as inventory.

The circuit can be viewed as a noninverting amplifier that amplifies V_p by the amount

$$A = \frac{V_o}{V_p} = 1 + \frac{R_2}{R_1} \tag{10.4}$$

where we are assuming an ideal op amp for simplicity. In turn, V_p is supplied by the op amp itself via the two *RC* networks as $V_p = [Z_p/(Z_p + Z_s)]V_o$, where $Z_p = R \parallel [1/(j2\pi fC)]$ and $Z_s = R + 1/(j2\pi fC)$. Expanding, we get

$$B(jf) = \frac{V_p}{V_o} = \frac{1}{3 + j(f/f_0 - f_0/f)} \tag{10.5}$$

where $f_0 = 1/(2\pi RC)$. The overall gain experienced by a signal in going around the loop is $T(jf) = AB$, or

$$T(jf) = \frac{1 + R_2/R_1}{3 + j(f/f_0 - f_0/f)} \tag{10.6}$$

This is a band-pass function since it approaches zero at both high and low frequencies. Its peak value occurs at $f = f_0$ and is

$$T(jf_0) = \frac{1 + R_2/R_1}{3} \tag{10.7}$$

The fact that $T(jf_0)$ is real indicates that a signal of frequency f_0 will experience a net phase shift of zero in going around the loop. Depending on the magnitude of $T(jf_0)$, we have three distinct possibilities:

1. $T(jf_0) < 1$, that is, $A < 3$ V/V. Any disturbance of frequency f_0 arising at the input of the op amp is first amplified by $A < 3$ V/V, and then by $B(jf_0) = \frac{1}{3}$ V/V, for a net gain of less than unity. Intuition tells us that this disturbance lessens each time it goes around the loop until it eventually decays to zero. We can state that negative feedback (via R_2 and R_1) prevails over positive feedback (via Z_s and Z_p), resulting in a stable system. Consequently, the circuit poles lie in the left half of the complex plane.
2. $T(jf_0) > 1$, that is, $A > 3$ V/V. Now positive feedback prevails over negative feedback, indicating that a disturbance of frequency f_0 will be amplified regeneratively, causing the circuit to break out into oscillations of growing magnitude. The circuit is now *unstable,* and its poles lie in the right half of the complex plane. As we know, the oscillations build up until the saturation limits of the op amp are reached. Thereafter, v_O will appear as a clipped sine wave when observed with the oscilloscope or visualized via PSpice.
3. $T(jf_0) = 1$, or $A = 3$ V/V *exactly,* a condition referred to as *neutral stability* because positive and negative feedback are now applied in equal amounts. Any disturbance of frequency f_0 is first amplified by 3 V/V and then by $\frac{1}{3}$ V/V, indicating that once started, it will be sustained indefinitely. As we know, this corresponds to a pole pair right on the $j\omega$ axis. The conditions $\measuredangle T(jf_0) = 0°$ and $|T(jf_0)| = 1$ are together referred to as the *Barkhausen criterion* for oscillation at $f = f_0$. The band-pass nature of $T(jf)$ allows for oscillation to occur only at $f = f_0$; any attempt to oscillate at other frequencies is naturally discouraged because $\measuredangle T \neq 0°$ and $|T| < 1$ there. By Eq. (10.7), neutral stability is achieved with

$$\frac{R_2}{R_1} = 2 \tag{10.8}$$

It is apparent that when this condition is met, the components around the op amp form a *balanced bridge* at $f = f_0$.

In a real-life circuit, component drift makes it difficult to keep the bridge exactly balanced. Moreover, provisions must be made so that (*a*) oscillation starts spontaneously at power turn-on, and (*b*) its amplitude is kept below the op amp saturation limits to avoid excessive distortion. These objectives are met by making the ratio R_2/R_1 *amplitude-dependent* such that at low signal levels it is slightly greater than 2 to ensure oscillation start-up, and that at high signal levels it is slightly less than 2 to limit amplitude. Then, once the oscillation has started, it will grow and automatically stabilize at some intermediate level where $R_2/R_1 = 2$ *exactly.*

Amplitude stabilization takes on many forms, all of which use nonlinear elements to either decrease R_2 or increase R_1 with signal amplitude. To provide an intuitive basis for our discussion, we shall continue using the function $T(jf)$, but in an incremental sense because of the nonlinearity now present in the circuit.

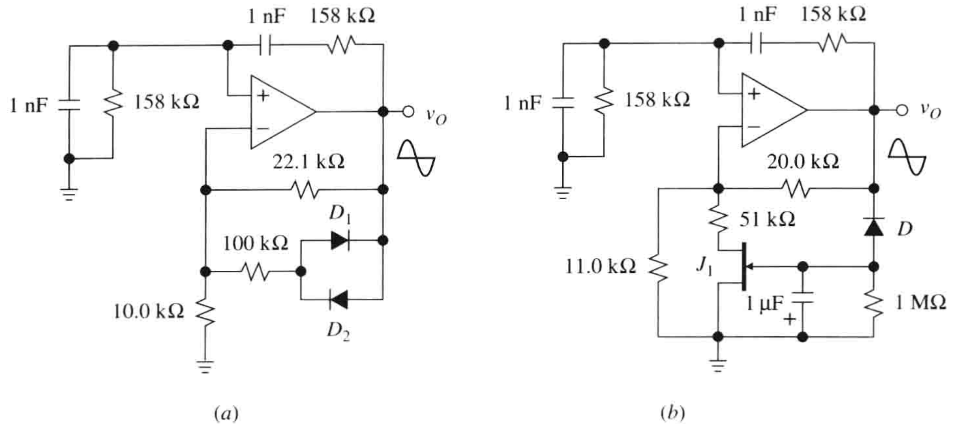


FIGURE 10.3
Practical Wien-bridge oscillators.

Automatic Amplitude Control

The circuit of Fig. 10.3*a* uses a simple diode-resistor network to control the effective value of R_2. At low signal levels the diodes are off, so the 100-kΩ resistance has no effect. We thus have $R_2/R_1 = 22.1/10.0 = 2.21$, or $T(jf_0) = (1 + 2.21)/3 = 1.07 > 1$, indicating oscillation buildup. As the oscillation grows, the diodes are gradually brought into conduction on alternate half-cycles. In the limit of heavy diode conduction, R_2 would effectively change to $(22.1 \parallel 100) = 18.1$ kΩ, giving $T(jf_0) = 0.937 < 1$. However, before this limiting condition is reached, amplitude will automatically stabilize at some intermediate level of diode conduction where $R_2/R_1 = 2$ exactly, or $T(jf_0) = 1$. We can easily visualize the process via the PSpice circuit of Fig. 10.4*a*, designed for an oscillation frequency of $f_0 = 1/[2\pi(159 \text{ k}\Omega) \times (1 \text{ nF})] \cong 1$ kHz. As shown in Fig. 10.4*b*, the oscillation starts up spontaneously and grows until its amplitude stabilizes at about 1 V.

A disadvantage of the above circuit is that V_{om} is quite sensitive to variations in the diode-forward voltage drops. The circuit of Fig. 10.3*b* overcomes this drawback by using an *n*-JFET as the stabilizing element.[1] At power turn-on, when the 1-μF capacitance is still discharged, the gate voltage is near 0 V, indicating a low channel resistance. The JFET effectively shorts the 51-kΩ resistance to ground to give $R_2/R_1 \cong 20.0/(11.0 \parallel 51) \cong 2.21 > 2$, so oscillation starts to build up. The diode and the 1-μF capacitance form a negative peak detector whose voltage becomes progressively more negative as the oscillation grows. This gradually reduces the conductivity of the JFET until, in the limit of complete cutoff we would have $R_2/R_1 = 20.0/11.0 = 1.82 < 2$. However, amplitude stabilizes automatically at some intermediate level where $R_2/R_1 = 2$ exactly. Denoting the corresponding gate-source voltage as $V_{GS(\text{crit})}$, and the output peak amplitude as V_{om}, we have $-V_{om} = V_{GS(\text{crit})} - V_{D(\text{on})}$. For instance, with $V_{GS(\text{crit})} = -4.3$ V we get $V_{om} \cong 4.3 + 0.7 = 5$ V.

Figure 10.5 shows yet another popular amplitude-stabilization scheme,[2] this time using a diode limiter for easier programming of amplitude. As usual, for low output levels the diodes are biased in cutoff, yielding $R_2/R_1 = 2.21 > 2$. The

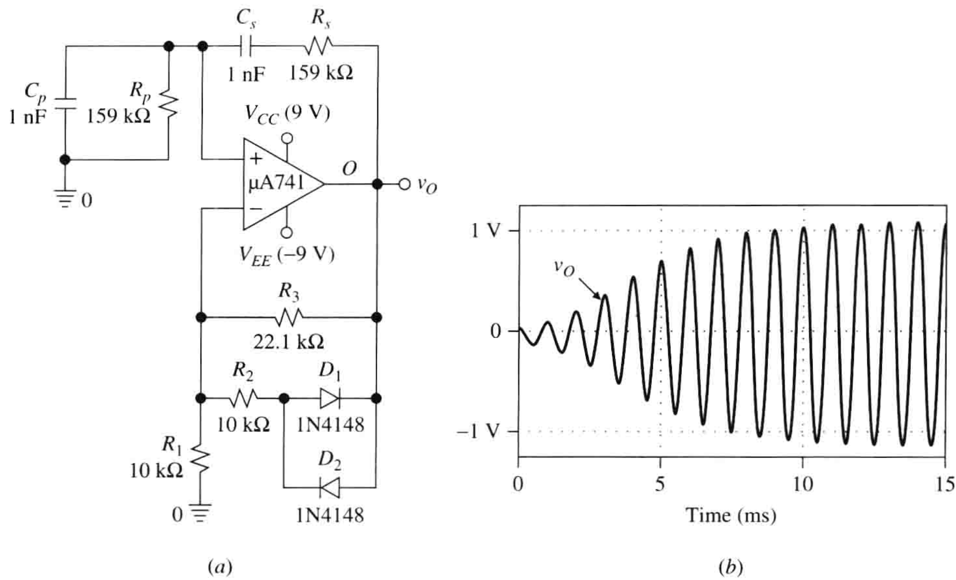


FIGURE 10.4
(a) Wien-bridge oscillator and (b) its output v_O.

oscillation grows until the diodes become conductive on alternate output peaks. Thanks to the symmetry of the clamping network, these peaks are likewise symmetric, or $\pm V_{om}$. To estimate V_{om}, consider the instant when D_2 starts to conduct. Assuming the current through D_2 is still negligible, and denoting the voltage at the anode of D_2 as V_2, we use KCL to write $(V_{om} - V_2)/R_3 \cong [V_2 - (-V_S)]/R_4$, where $V_2 = V_n + V_{D2(\text{on})} \cong V_{om}/3 + V_{D2(\text{on})}$. Eliminating V_2 and solving gives $V_{om} \cong 3[(1 + R_4/R_3)V_{D2(\text{on})} + V_S]/(2R_4/R_3 - 1)$. For example, with $R_3 = 3\ \text{k}\Omega$, $R_4 = 20\ \text{k}\Omega$, $V_S = 15$ V, and $V_{D(\text{on})} = 0.7$ V, we get $V_{om} \cong 5$ V.

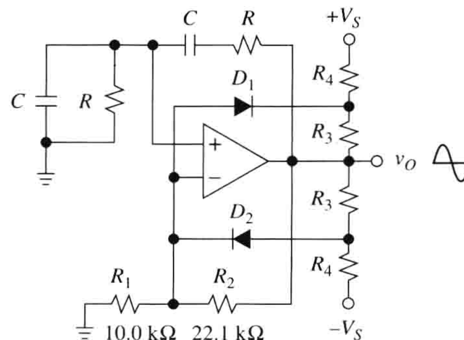


FIGURE 10.5
Wien-bridge oscillator using a limiter for amplitude stabilization.

Practical Considerations

The accuracy and stability of the oscillation are affected by the quality of the passive components as well as op amp dynamics. Good choices for the elements in the positive-feedback network are polycarbonate capacitors and thin-film resistors. To compensate for component tolerances, practical Wien-bridge circuits are often equipped with suitable trimmers for the exact adjustment of f_0 as well as THD minimization. With proper trimming, THD levels as low as 0.01% can be achieved.[1] We observe that because of the filtering action provided by the positive-feedback network, the sine wave v_P available at the noninverting input is generally purer than v_O. Consequently, it may be desirable to use v_P as the output, though a buffer would be needed to avoid perturbing circuit behavior.

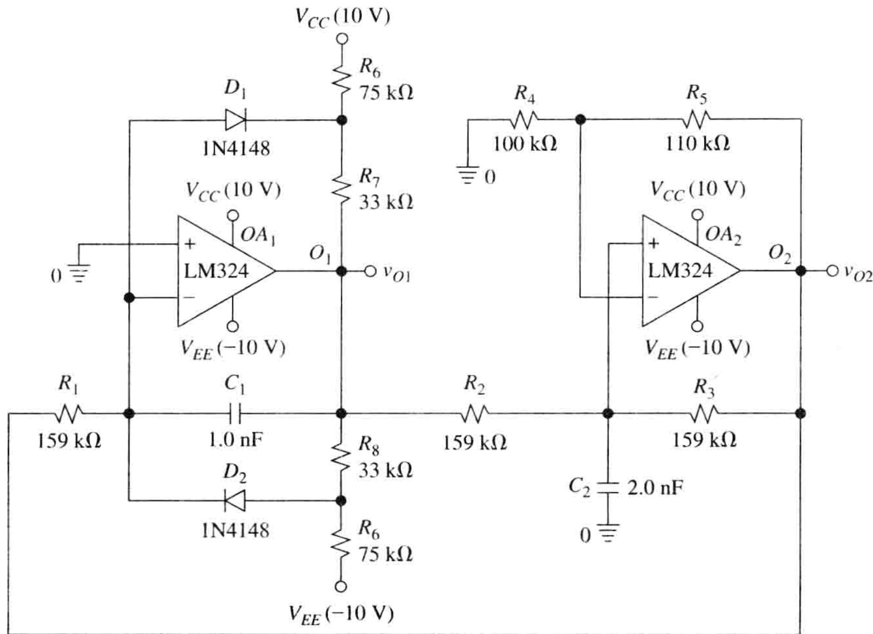
To avoid slew-rate limiting effects for a given output peak-amplitude V_{om}, the op amp should have $\text{SR} > 2\pi V_{om} f_0$. Once this condition is met, the limiting factor becomes the finite GBP, whose effect is a downshift in the actual frequency of oscillation. It can be proved[2] that to contain this shift within 10% when a constant-GBP op amp is used, the latter should have $\text{GBP} \geq 43 f_0$. To compensate for this downshift, one can suitably predistort the element values of the positive-feedback network, in a manner similar to the filter predistortion techniques of Sections 6.5 and 6.6.

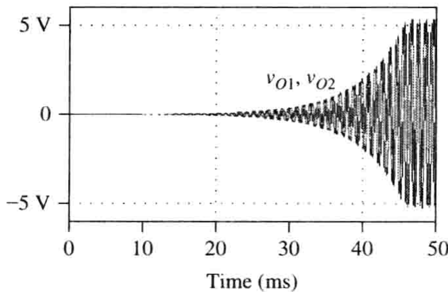
The low end of the frequency range depends on how large the components in the reactive network can be made. Using FET-input op amps to minimize input-bias-current errors, the value of R can easily be increased to the range of tens of megohms. For instance, using $C = 1\ \mu\text{F}$ and $R = 15.9\ \text{M}\Omega$ gives $f_0 = 0.01$ Hz.

Quadrature Oscillators

We can generalize the above ideas and make an oscillator out of any second-order filter that is capable of giving $Q = \infty$ as well as $Q < 0$. To this end, we first ground the input, since it is no longer necessary; then, we design for an initially negative Q to force the poles in the right half of the complex plane and thus ensure oscillation startup; finally, we include a suitable amplitude-dependent network to automatically pull the poles back to the $j\omega$ axis and give $Q = \infty$, or sustained oscillation.

Of special interest are filter topologies of the dual-integrator-loop type, since they provide two oscillations in quadrature, that is, with a relative phase shift of 90°. A good candidate is the loop of Fig. 10.6*a*, consisting of the inverting integrator OA_1 and the noninverting (or Deboo) integrator OA_2. (The circuit uses half the LM334 quad op amp, whose macromodel is available in PSpice's library.) With the component values shown, both integrators admit, ideally, a unity-gain frequency of $f_0 = 1/[2\pi(159\ \text{k}\Omega) \times (1\ \text{nF})] \cong 1$ kHz. In practice, because of the open-loop gain rolloff of the op amps, the integrators will be slightly lossy, as discussed in Chapter 6. To compensate for these losses as well as to make the circuit start out with $Q < 0$, we impose $R_5 > R_4$, as per Eq. (3.22). Under this condition, the oscillation will build up as in Fig. 10.6*b*, until the diode network discussed in connection with Fig. 10.5 takes over and stabilizes the amplitude of the oscillation at about 5 V. As depicted in the expanded view of Fig. 10.6*c*, v_{O1} exhibits some distortion, but v_{O2} is purer because of the filtering action provided by the Deboo integrator.





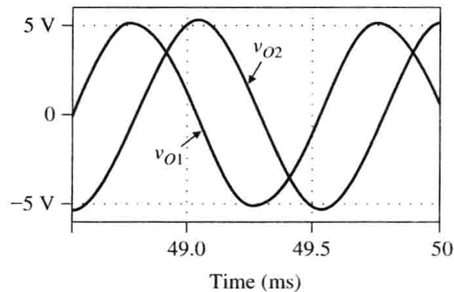


FIGURE 10.6
(*a*) Quadrature oscillator, (*b*) growing oscillations, and (*c*) expanded view after amplitude stabilization.

10.2 MULTIVIBRATORS

Multivibrators are regenerative circuits intended especially for timing applications. Multivibrators are classified as bistable, astable, and monostable.

In a *bistable multivibrator* both states are stable, so external commands are needed to force the circuit to a given state. This is the popular *flip-flop,* which in turn takes on different names, depending on the way in which the external commands are effected.

An *astable multivibrator* toggles spontaneously between one state and the other, without any external commands. Also called a *free-running multivibrator,* its timing is set by a suitable network, usually comprising a capacitor or a quartz crystal.

A *monostable multivibrator,* also called a *one-shot,* is stable only in one of its two states. If forced into the other state via an external command called a *trigger,* it returns to its stable state spontaneously, after a delay set by a suitable timing network.

Here we are interested in astable and monostable multivibrators. These circuits are implemented with voltage comparators or with logic gates, especially CMOS gates.

Basic Free-Running Multivibrator

In the circuit of Fig. 10.7*a*, the 301 op amp comparator and the positive-feedback resistances R_1 and R_2 form an inverting Schmitt trigger. Assuming symmetric output saturation at $\pm V_{\text{sat}} = \pm 13$ V, the Schmitt-trigger thresholds are also symmetric at $\pm V_T = \pm V_{\text{sat}} R_1/(R_1 + R_2) = \pm 5$ V. The signal to the inverting input is provided by the op amp itself via the RC network.

At power turn-on ($t = 0$) v_O will swing either to $+V_{\text{sat}}$ or to $-V_{\text{sat}}$, since these are the only stable states admitted by the Schmitt trigger. Assume it swings to $+V_{\text{sat}}$, so that $v_P = +V_T$. This will cause R to charge C toward V_{sat}, leading to an exponential rise in v_N with the time constant $\tau = RC$. As soon as v_N catches up with $v_P = V_T$, v_O snaps to $-V_{\text{sat}}$, reversing the capacitance current and also causing v_P to snap to $-V_T$. So, now v_N decays exponentially toward $-V_{\text{sat}}$ until it catches up with $v_P = -V_T$, at which point v_O again snaps to $+V_{\text{sat}}$, thus repeating the cycle. It is apparent that once powered, the circuit has the ability to start and then sustain oscillation, with v_O snapping back and forth between $+V_{\text{sat}}$ and $-V_{\text{sat}}$, and v_N slewing exponentially back and forth between $+V_T$ and $-V_T$. After the power-on cycle, the waveforms become periodic.

We are interested in the frequency of oscillation, which is found from the period T as $f_0 = 1/T$. Thanks to the symmetry of the saturation levels, v_O has a duty cycle of 50%, so we only need to find $T/2$. Applying Eq. (10.3) with $\Delta t = T/2$, $\tau = RC$,

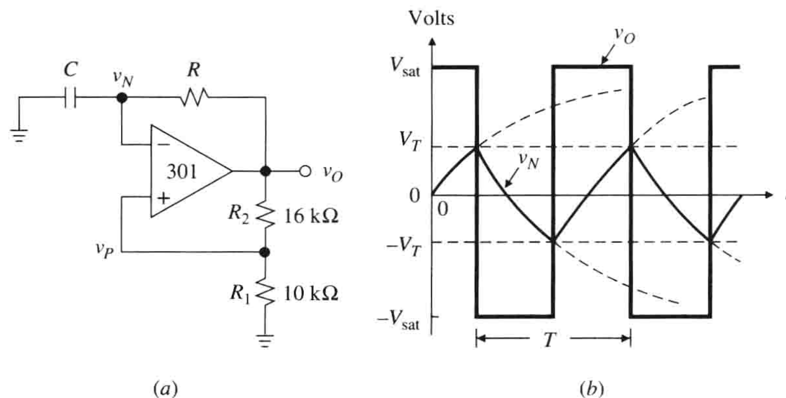


FIGURE 10.7
Basic free-running multivibrator.

$V_\infty = V_{sat}$, $V_0 = -V_T$, and $V_1 = +V_T$, we get

$$\frac{T}{2} = RC \ln \frac{V_{sat} + V_T}{V_{sat} - V_T}$$

Substituting $V_T = V_{sat}/(1 + R_2/R_1)$ and simplifying finally gives

$$f_0 = \frac{1}{T} = \frac{1}{2RC \ln(1 + 2R_1/R_2)} \tag{10.9}$$

With the components shown, $f_0 = 1/(1.62RC)$. If we use the ratio $R_1/R_2 = 0.859$, then $f_0 = 1/(2RC)$.

We observe that f_0 depends only on the external components. In particular, it is unaffected by V_{sat}, which is known to be an ill-defined parameter since it varies from one op amp to another and also depends on the supply voltages. Any variation in V_{sat} will cause V_T to vary in proportion, thus ensuring the same transition time and, hence, the same oscillation frequency.

The maximum operating frequency is determined by the comparator speed. With the 301 op amp as a comparator, the circuit yields a reasonably good square wave up to the 10-kHz range. This can be extended significantly by using a faster device. At higher frequencies, however, the stray capacitance of the noninverting input toward ground becomes a limiting factor. This can be compensated by using a suitable capacitance in parallel with R_2.

The lowest operating frequency depends on the practical upper limits of R and C, as well as the net leakage at the inverting input node. FET-input comparators may be a good choice in this case.

Although f_0 is unaffected by uncertainties in V_{sat}, it is often desirable to stabilize the output levels for a cleaner and more predictable square-wave amplitude. This is readily achieved with a suitable voltage-clamping network. If it is desired to vary f_0, a convenient approach is to use an array of decade capacitances and a rotary switch for decade selection, and a variable resistance for continuous tuning within the selected decade.

EXAMPLE 10.1. Design a square-wave generator meeting the following specifications: (*a*) f_0 must be variable in decade steps from 1 Hz to 10 kHz; (*b*) f_0 must be variable continuously within each decade interval; (*c*) amplitude must be ± 5 V, stabilized. Assume ± 15-V poorly regulated supplies.

Solution. To ensure stable ± 5 V output levels, use a diode-bridge clamp as in Fig. 10.8. When the op amp saturates at $+13$ V, current flows through the path R_3-D_1-D_5-D_4, thus clamping v_O at $V_{D1(on)} + V_{Z5} + V_{D4(on)}$. To clamp at 5 V, use $V_{Z5} = 5 - 2V_{D(on)} = 5 - 2 \times 0.7 = 3.6$ V. When the op amp saturates at -13 V, current flows through the path D_3-D_5-D_2-R_3, clamping v_O at -5 V.

To vary f_0 in decade steps, use the four capacitances and rotary switch shown. To vary f_0 within a given decade, implement R with a pot. To cope with component tolerances, ensure an adequate amount of overlap between adjacent decade intervals. To be on the safe side, impose a range of continuous variability from 0.5 to 20, that is, over a 40-to-1 range. We then have $R_{pot} + R_s = 40R_s$, or $R_{pot} = 39R_s$. To keep input-bias-current errors low, impose $I_{R(min)} \gg I_B$, say, $I_{R(min)} = 10\ \mu$A. Moreover, let $R_1 = R_2 = 33$ kΩ, so that $V_T = 2.5$ V. Then, $R_{max} = (5 - 2.5)/(10 \times 10^{-6}) = 250$ kΩ. Since $R_s \ll R_{pot}$, use a 250-kΩ pot. Then, $R_s = 250/39 = 6.4$ kΩ (use 6.2 kΩ).

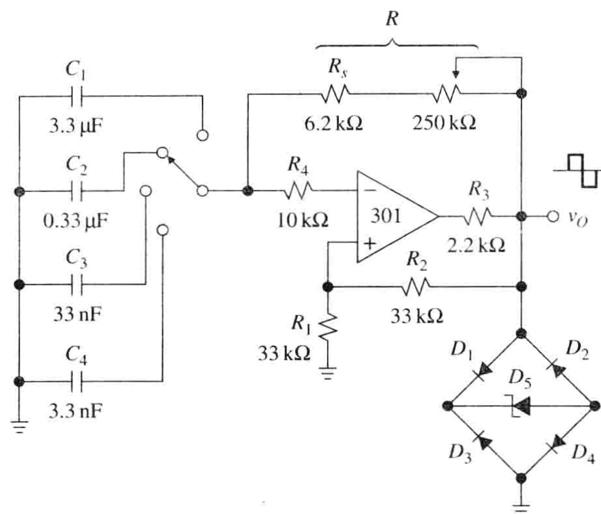


FIGURE 10.8
Square-wave generator of Example 10.1.

To find C_1, impose $f_0 = 0.5$ Hz with the pot set to its maximum value. By Eq. (10.9), $C_1 = 1/[2 \times 0.5 \times (250 + 6.2) \times 10^3 \times \ln 3] = 3.47\ \mu$F. The closest standard value is $C_1 = 3.3\ \mu$F. Then, $C_2 = 0.33\ \mu$F, $C_3 = 33$ nF, and $C_4 = 3.3$ nF.

The function of R_4 is to protect the comparator input stage at power turn-off, when the capacitors may still be charged, and that of R_3 is to supply current to the bridge, R_2, R, and to the external load, if any. The maximum current drawn by R is when $v_O = +5$ V, $v_N = -2.5$ V, and the pot is set to zero. This current is $[5 - (-2.5)]/6.2 = 1.2$ mA. We also have $I_{R_2} = 5/66 = 0.07$ mA. Imposing a bridge current of 1 mA and allowing for a maximum load current of 1 mA, we have $I_{R_3(\max)} = 1.2 + 0.07 + 1 + 1 \cong 3.3$ mA. Hence, $R_3 = (13 - 5)/3.3 = 2.4$ kΩ (use 2.2 kΩ to be safe). For the diode bridge use a CA3039 array.

Figure 10.9 shows a multivibrator designed for single-supply operation. By using a fast comparator, the circuit can operate well into the hundreds of kilohertz. As we know, the circuit gives $V_{OL} \cong 0$ and, if $R_4 \ll R_3 + (R_1 \parallel R_2)$, it gives $V_{OH} \cong V_{CC}$. At power turn-on ($t = 0$), when C is still discharged, v_O is forced high, causing C to charge toward V_{CC} via R. As soon as v_N reaches V_{TH}, v_O snaps low, causing C to discharge toward ground. Henceforth, the oscillation becomes periodic with duty cycle $D(\%) = 100T_H/(T_L + T_H)$ and $f_0 = 1/(T_L + T_H)$. Applying Eq. (10.3) twice, first with $\Delta t = T_L$, $V_\infty = 0$, $V_0 = V_{TH}$, and $V_1 = V_{TL}$, then with $\Delta t = T_H$, $V_\infty = V_{CC}$, $V_0 = V_{TL}$, and $V_1 = V_{TH}$, we get, after combining terms,

$$f_0 = \frac{1}{RC \ln \left(\dfrac{V_{TH}}{V_{TL}} \times \dfrac{V_{CC} - V_{TL}}{V_{CC} - V_{TH}} \right)} \tag{10.10}$$

To simply inventory and achieve $D = 50\%$, it is customary to impose $R_1 = R_2 = R_3$, after which $f_0 = 1/(RC \ln 4) = 1/1.39RC$. Oscillators of this type can easily achieve stabilities approaching 0.1% with initial predictability of the order of 5% to 10%.

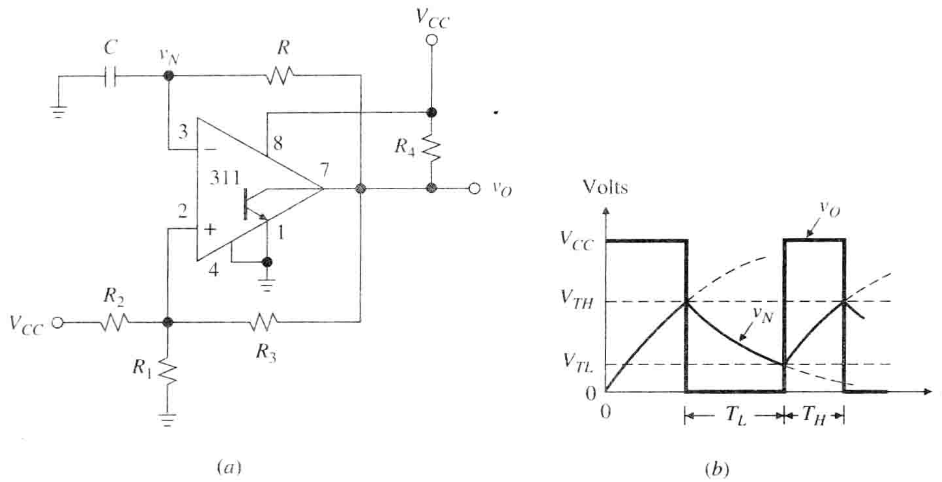


FIGURE 10.9
Single-supply free-running multivibrator.

EXAMPLE 10.2. In the circuit of Fig. 10.9 specify components for $f_0 = 1$ kHz, and verify with PSpice for $V_{CC} = 5$ V.

Solution. Use $R_1 = R_2 = R_3 = 33\ \text{k}\Omega$, $R_4 = 2.2\ \text{k}\Omega$, $C = 10$ nF, and $R = 73.2\ \text{k}\Omega$. Using the PSpice circuit of Fig. 10.10*a*, we readily obtain the waveforms of Fig. 10.10*b*. Cursor measurements give $T = 1.002$ ms, or $f_0 = 998$ Hz.

Free-Running Multivibrator Using CMOS Gates

CMOS logic gates are particularly attractive when analog and digital functions must coexist on the same chip. A CMOS gate enjoys an extremely high input impedance, a rail-to-rail input range and output swing, extremely low power consumption, and

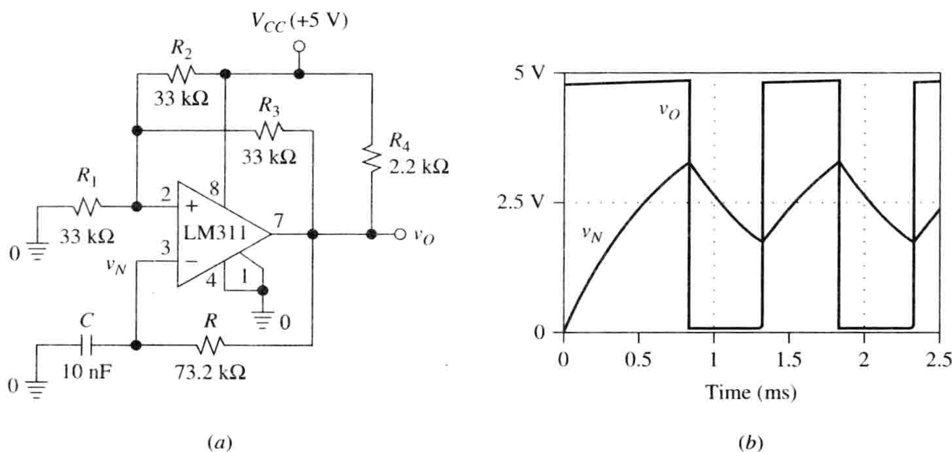


FIGURE 10.10
(*a*) Free-running multivibrator of Example 10.2, and (*b*) waveforms.

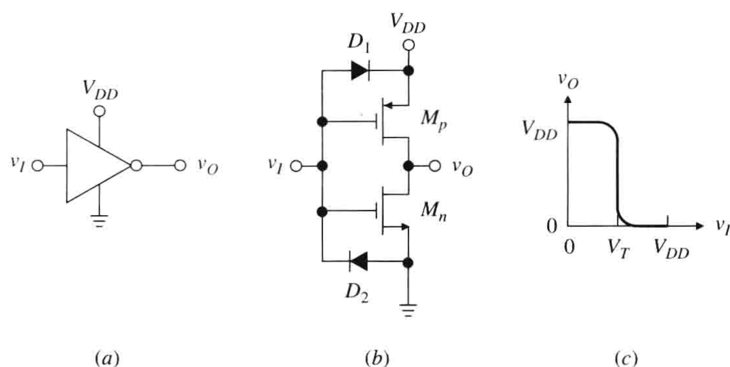


FIGURE 10.11
CMOS inverter: logic symbol, internal circuit diagram, and VTC.

the speed and low cost of logic circuitry. The simplest gate is the inverter depicted in Fig. 10.11. This gate can be regarded as an inverting-type threshold detector giving $v_O = V_{OH} = V_{DD}$ for $v_I < V_T$, and $v_O = V_{OL} = 0$ for $v_I > V_T$. The threshold V_T is the result of internal transistor operation, and it is nominally halfway between V_{DD} and 0, or $V_T \cong V_{DD}/2$. The protective diodes, normally in cutoff, prevent v_I from rising above $V_{DD} + V_{D(\text{on})}$ or dropping below $-V_{D(\text{on})}$, and thus protect the FETs against possible electrostatic discharge.

In the circuit of Fig. 10.12*a* assume at power turn-on ($t = 0$) v_2 goes high. Then, by I_2's inverting action, v_O remains low, and C starts charging toward $v_2 = V_{DD}$ via R. The ensuing exponential rise is conveyed to I_1 via R_1 as signal v_1. As soon as

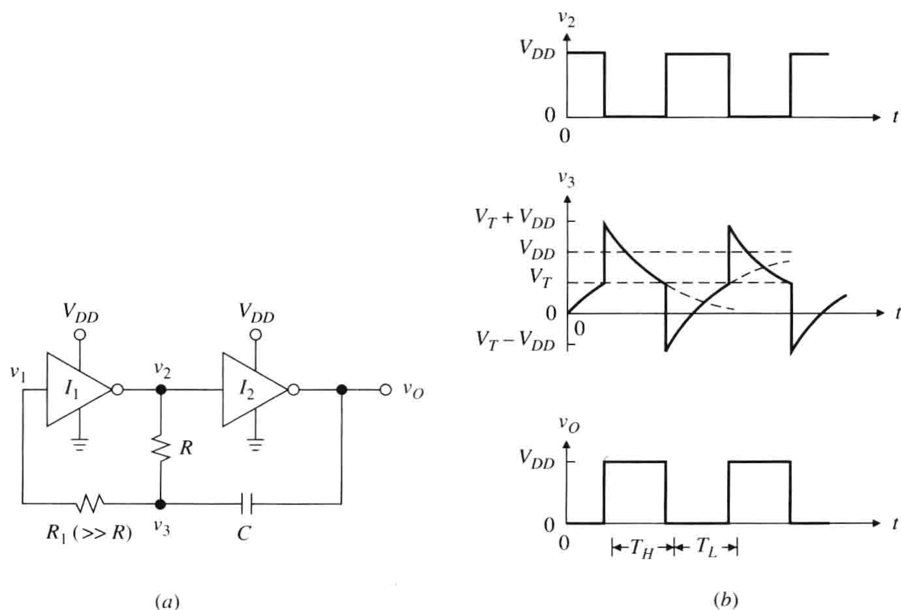


FIGURE 10.12
CMOS-gate free-running multivibrator.

v_1 rises to V_T, I_1 changes state and pulls v_2 low, forcing I_2 to pull v_O high. Since the voltage across C cannot change instantaneously, the step change in v_O causes v_3 to change from V_T to $V_T + V_{DD} \cong 1.5V_{DD}$, as shown in the timing diagram. These changes occur by a snapping action similar to that of Schmitt triggers.

With v_3 being high and v_2 being low, C will now discharge toward $v_2 = 0$ via R. As soon as the value of v_3 decays to V_T, the circuit snaps back to the previous state; that is, v_2 goes high and v_O goes low. The step change in v_O causes v_3 to jump from V_T to $V_T - V_{DD} \cong -0.5V_{DD}$, after which v_3 will again charge toward $v_2 = V_{DD}$. As shown, v_2 and v_O snap back and forth between 0 and V_{DD}, but in antiphase, and they snap each time v_3 reaches V_T.

To find $f_0 = 1/(T_H + T_L)$, we again use Eq. (10.3), first with $\Delta t = T_H$, $V_\infty = 0$, $V_0 = V_T + V_{DD}$, and $V_1 = V_T$, then with $\Delta t = T_L$, $V_\infty = V_{DD}$, $V_0 = V_T - V_{DD}$, and $V_1 = V_T$. The result is

$$f_0 = \frac{1}{RC \ln\left(\dfrac{V_{DD} + V_T}{V_T} \times \dfrac{2V_{DD} - V_T}{V_{DD} - V_T}\right)} \tag{10.11}$$

For $V_T = V_{DD}/2$ we get $f_0 = 1/(RC \ln 9) = 1/2.2RC$ and $D(\%) = 50\%$. In practice, due to production variations, there is a spread in the values of V_T. This, in turn, affects f_0, thus limiting the circuit to applications where frequency accuracy is not of primary concern.

We observe that if v_3 were applied to I_1 directly, the input protective diodes of I_1 would clamp v_3 and alter the timing significantly. This is avoided by using the decoupling resistance $R_1 \gg R$ (in practice, $R_1 \cong 10R$ will suffice).

CMOS Crystal Oscillator

In precise timekeeping applications, frequency must be much more accurate and stable than that afforded by simple RC oscillators. These demands are met with crystal oscillators, an example of which is shown in Fig. 10.13. Since the circuit exploits the electromechanical-resonance characteristics[3] of a quartz crystal to set f_0, it acts more like a tuned amplifier than a multivibrator. The idea here is to

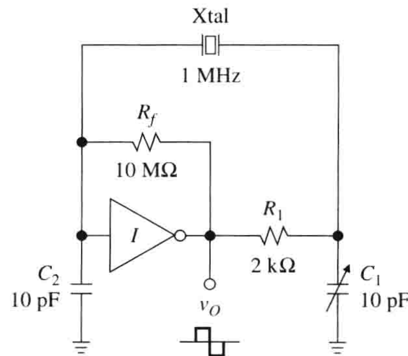


FIGURE 10.13
CMOS-gate crystal oscillator.

place a network that includes a crystal in the feedback loop of a high-gain inverting amplifier. This network routes a portion of the output signal back to the input, where it is reamplified in such a way as to sustain oscillation at a frequency set by the crystal.

A CMOS gate is made to operate as a high-gain amplifier by biasing it near the center of its VTC, where slope is the steepest and gain is thus maximized. Using a plain feedback resistance R_f, as shown, establishes the dc operating point at $V_O = V_I = V_T \cong V_{DD}/2$. Thanks to the extremely low input leakage current of CMOS gates, R_f can be made quite large. The function of the remaining components is to help establish the proper loss and phase, as well as provide a low-pass filter action to discourage oscillation at the crystal's higher harmonics.

Although crystals have to be ordered for specific frequencies, a number of commonly used units are available off the shelf, namely, 32.768 kHz crystals for digital wristwatches, 3.579545 MHz for TV tuners, and 100 kHz, 1 MHz, 2 MHz, 4 MHz, 5 MHz, 10 MHz, etc., for digital clock applications. A crystal oscillator can be tuned slightly by varying one of its capacitances, as shown. Crystal oscillators of the type shown can easily achieve stabilities on the order of 1 ppm/°C (1 part-per-million per degree Celsius).[4]

The duty cycle of clock generators is not necessarily 50%. Applications requiring perfect square-wave symmetry are easily accommodated by feeding the oscillator to a toggle flip-flop. The latter then produces a square wave with $D(\%) = 50\%$, but with half the frequency of the oscillator. To achieve the desired frequency, we simply use a crystal with a frequency rating twice as high.

Monostable Multivibrator

On receiving a trigger pulse at the input, a monostable multivibrator or one-shot produces a pulse of a specified duration T. This duration can be generated digitally, by counting a specified number of pulses from a clock source, or in analog fashion, by using a capacitor for time-out control. One-shots are used to generate strobe commands and delays, and in switch debouncing.

The circuit of Fig. 10.14 uses a NOR gate G and an inverter I. The NOR yields a high output only when both inputs are low; if at least one of the inputs is high, the output will be low. Under normal conditions, v_I is low and C is in steady state, so $v_2 = V_{DD}$ due to the pullup action by R, and $v_O = 0$ by inverter action. Further, since both inputs to the NOR gate are low, its output is high, or $v_1 = V_{DD}$, indicating zero voltage across C.

The arrival of a trigger pulse v_I causes the NOR gate to pull v_1 low. Since the voltage across C cannot change instantaneously, v_2 will also go low, causing in turn v_O to go high. Even if the trigger pulse is now deactivated, the NOR gate will keep v_1 low because v_O is high. This state of affairs, however, cannot last indefinitely because R is now charging C toward V_{DD}. In fact, as soon as v_2 reaches V_T, the inverter snaps, forcing v_O back low. In response to this, the NOR gate forces v_1 high, and C then transmits this step to the inverter, thus reinforcing its initial snap in Schmitt-trigger fashion. Even though v_2 tries to swing from V_T to $V_T + V_{DD} \cong 1.5V_{DD}$, the internal protective diode D_1 of the inverter, shown explicitly in Fig. 10.11b, will clamp v_2 near V_{DD}, thus discharging C. The circuit is now back in the stable state preceding

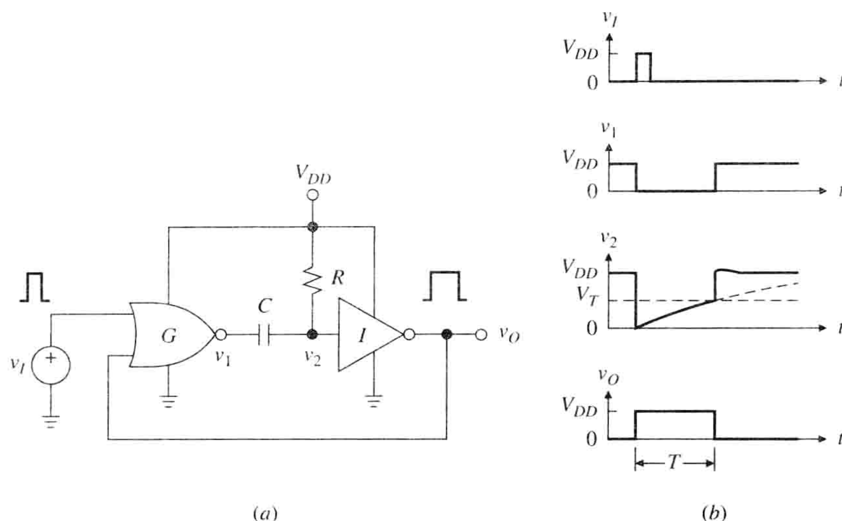


FIGURE 10.14
CMOS-gate one-shot.

the arrival of the trigger pulse. The timeout T is found via Eq. (10.3) as

$$T = RC \ln \frac{V_{DD}}{V_{DD} - V_T} \tag{10.12}$$

For $V_T = V_{DD}/2$, this reduces to $T = RC \ln 2 = 0.69RC$.

A *retriggerable* one-shot begins a new cycle each time the trigger is activated, including activation during T. By contrast, a *nonretriggerable* one-shot is insensitive to triggering during T.

10.3 MONOLITHIC TIMERS

The need for the astable and monostable functions arises so often that special circuits,[4] called *IC timers*, are available to satisfy these needs. Among the variety of available products, the one that has gained the widest acceptance in terms of cost and versatility is the 555 timer. Another popular product is the 2240 timer, which combines a timer with a programmable counter to provide additional timing flexibility.

The 555 Timer

As shown in Fig. 10.15, the basic blocks of the 555 timer are: (*a*) a trio of identical resistors, (*b*) a pair of voltage comparators, (*c*) a flip-flop, and (*d*) a BJT switch Q_O. The resistances set the comparator thresholds at $V_{TH} = (2/3)V_{CC}$ and $V_{TL} = (1/3)V_{CC}$. For additional flexibility, the upper threshold node is externally accessible via pin 5 so that the user can modulate the value of V_{TH}. Regardless of the value of V_{TH}, we always have $V_{TL} = V_{TH}/2$.

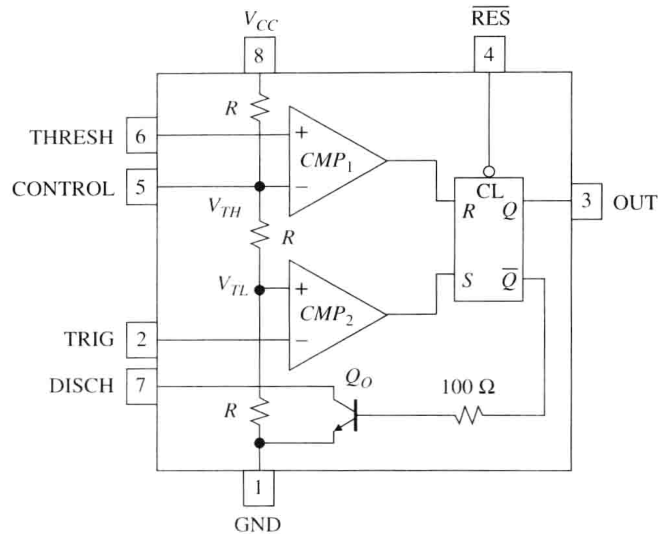


FIGURE 10.15
The 555 timer block diagram.

The state of the flip-flop is controlled by the comparators as follows: (*a*) Whenever the voltage at the trigger input (TRIG) drops below V_{TL}, CMP_2 fires and sets the flip-flop, forcing Q high and $\overline{Q}$ low; with a low voltage at its base, Q_O is in cutoff. (*b*) Whenever the voltage at the threshold input (THRESH) rises above V_{TH}, CMP_1 fires and clears the flip-flop, forcing Q low and $\overline{Q}$ high. With a high voltage applied to its base via the 100-Ω resistance, Q_O is now heavily on. Summarizing, lowering TRIG below V_{TL} turns Q_O off, and raising THRESH above V_{TH} turns Q_O heavily on. The flip-flop includes a reset input ($\overline{\text{RES}}$) to force Q low and turn Q_O on regardless of the conditions at the inputs of the comparators.

The 555 is available in both bipolar and CMOS versions. The bipolar versions operate over a wide range of supply voltages, typically 4.5 V $\leq V_{CC} \leq$ 18 V, and are capable of sourcing and sinking output currents of 200 mA. The TLC555, which is a popular CMOS version, is designed to operate over a power-supply range of 2 V to 18 V, and has output current sinking and sourcing capabilities of 100 mA and 10 mA, respectively. The transistor switch is an enhancement-type *n*-MOSFET. The advantages of CMOS timers are low power consumption, very high input impedances, and a rail-to-rail output swing.

The 555 as an Astable Multivibrator

Figure 10.16 shows how the 555 is configured for astable operation using just three external components. To understand circuit operation, refer also to the internal diagram of Fig. 10.15.

At power turn-on ($t = 0$), when the capacitor is still discharged, the voltage at the TRIG input is less than V_{TL}. This forces Q high and keeps the BJT in cutoff, thus allowing C to charge toward V_{CC} via the series $R_A + R_B$. As soon as v_C reaches V_{TH},

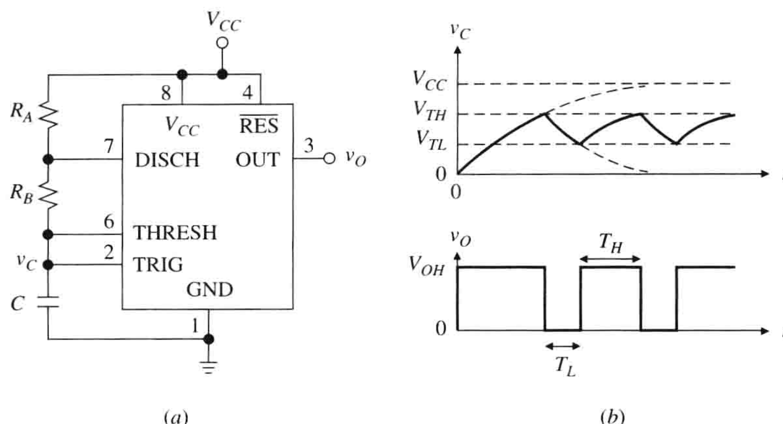


FIGURE 10.16
The 555 timer as an astable multivibrator.

CMP_1 fires and forces Q low. This turns on Q_O, which then pulls the DISCH pin to $V_{CE(\text{sat})} \cong 0$ V. Consequently, C now discharges toward ground via R_B. As soon as v_C reaches V_{TL}, CMP_2 fires, forcing Q high and turning off Q_O. This reestablishes the conditions for a new cycle of astable operation.

The time intervals T_L and T_H are found via Eq. (10.3). During T_L the time constant is R_BC, so $T_L = R_BC \ln(V_{TH}/V_{TL}) = R_BC \ln 2$; during T_H the time constant is $(R_A + R_B)C$, so $T_H = (R_A + R_B)C \ln[(V_{CC} - V_{TL})/(V_{CC} - V_{TH})]$. Consequently,

$$T = T_L + T_H = R_BC \ln 2 + (R_A + R_B)C \ln \frac{V_{CC} - V_{TH}/2}{V_{CC} - V_{TH}} \qquad (10.13)$$

Substituting $V_{TH} = (2/3)V_{CC}$ and solving for $f_0 = 1/T$ and $D(\%) = 100T_H/(T_L + T_H)$ gives

$$f_0 = \frac{1.44}{(R_A + 2R_B)C} \qquad D(\%) = 100\frac{R_A + R_B}{R_A + 2R_B} \qquad (10.14)$$

We observe that the oscillation characteristics are set by the external components and are independent of V_{CC}. To prevent power-supply noise from causing false triggering when v_C approaches either threshold, use a 0.01-μF bypass capacitor between pin 5 and ground: this will clean V_{TH} as well as V_{TL}. The timing accuracy[4] of the 555 astable approaches 1%, with a temperature stability of 0.005%/°C and a power-supply stability of 0.05%/V.

EXAMPLE 10.3. In the circuit of Fig. 10.16 specify suitable components for $f_0 = 50$ kHz and $D(\%) = 75\%$.

Solution. Let $C = 1$ nF, so that $R_A + 2R_B = 1.44/(f_0C) = 28.85$ kΩ. Imposing $(R_A + R_B)/(R_A + 2R_B) = 0.75$ gives $R_A = 2R_B$. Solving gives $R_A = 14.4$ kΩ (use 14.3 kΩ) and $R_B = 7.21$ kΩ (use 7.15 kΩ).

Since V_{TL} and V_{TH} remain stable during the oscillation cycle, the dual-comparator scheme utilized in the 555 allows higher operating frequencies than

the single-comparator schemes of the previous section. In fact some 555 versions can easily operate to the megahertz range. The upper frequency limit is determined by the combined propagation delays of the comparators, flip-flop, and transistor switch. The lower-frequency limit is determined by how large the external component values can practically be made. Thanks to the extremely low input currents, CMOS timers allow for large external resistances, so very long time constants can be obtained without using excessively large capacitances.

Since $T_H > T_L$, the circuit always gives $D(\%) > 50\%$. A symmetric duty cycle can be approached in the limit $R_A \ll R_B$; however, making R_A too small may lead to excessive power dissipation. A better approach to perfect symmetry is to use an output toggle flip-flop, as discussed in the previous section.

The 555 as a Monostable Multivibrator

Figure 10.17 shows the 555 connection for monostable operation. Under normal conditions, the TRIG input is held high, and the circuit is in the stable state represented by Q low. Moreover, the BJT switch Q_O is closed, keeping C discharged, or $v_C \cong 0$.

The circuit is triggered by lowering the TRIG input below V_{TL}. When this is done, CMP_2 sets the flip-flop, forcing Q high and turning off Q_O. This frees C to charge toward V_{CC} via R. However, as soon as v_C reaches V_{TH}, the upper comparator clears the flip-flop, forcing Q low and turning Q_O heavily on. The capacitance is rapidly discharged, and the circuit returns to the stable state preceding the arrival of the trigger pulse.

The pulse width T is readily found via Eq. (10.3) as

$$T = RC \ln \frac{V_{CC}}{V_{CC} - V_{TH}} \tag{10.15}$$

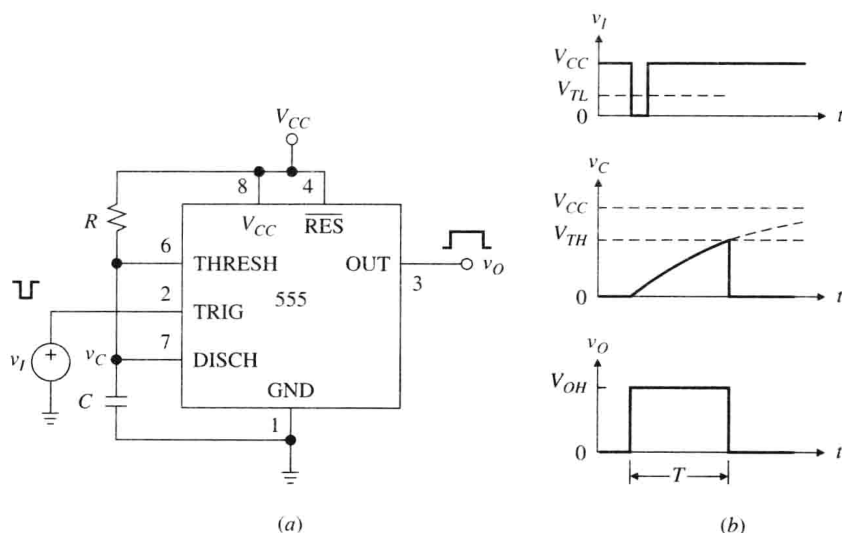


FIGURE 10.17
The 555 timer as a monostable multivibrator.

Letting $V_{TH} = (2/3)V_{CC}$ gives $T = RC \ln 3$, or

$$T = 1.10RC \tag{10.16}$$

Note once again the independence of V_{CC}. To enhance noise immunity, connect a 0.01-μF capacitor between pin 5 and ground (see Fig 10.15).

Voltage Control

If desired, the timing characteristics of the 555 can be modulated via the CONTROL input. Changing V_{TH} from its nominal value of $(2/3)V_{CC}$ will result in longer or shorter capacitance charging times, depending on whether V_{TH} is increased or decreased.

When the timer is configured for astable operation, modulating V_{TH} varies T_H while leaving T_L unchanged, as indicated by Eq. (10.13). Consequently, the output is a train of constant-width pulses with a variable repetition rate. This is referred to as *pulse-position modulation* (PPM).

When the timer is configured for monostable operation, modulating V_{TH} varies T, as per Eq. (10.15). If the monostable is triggered by a continuous pulse train, the output will be a pulse train with the same frequency as the input but with the pulse width modulated by V_{TH}. We now have *pulse-width modulation* (PWM).

PPM and PWM represent two common forms of information encoding for storage and transmission. Note that once V_{TH} is overridden externally, V_{TH} and V_{CC} are no longer related; hence, the timing characteristics are no longer independent of V_{CC}.

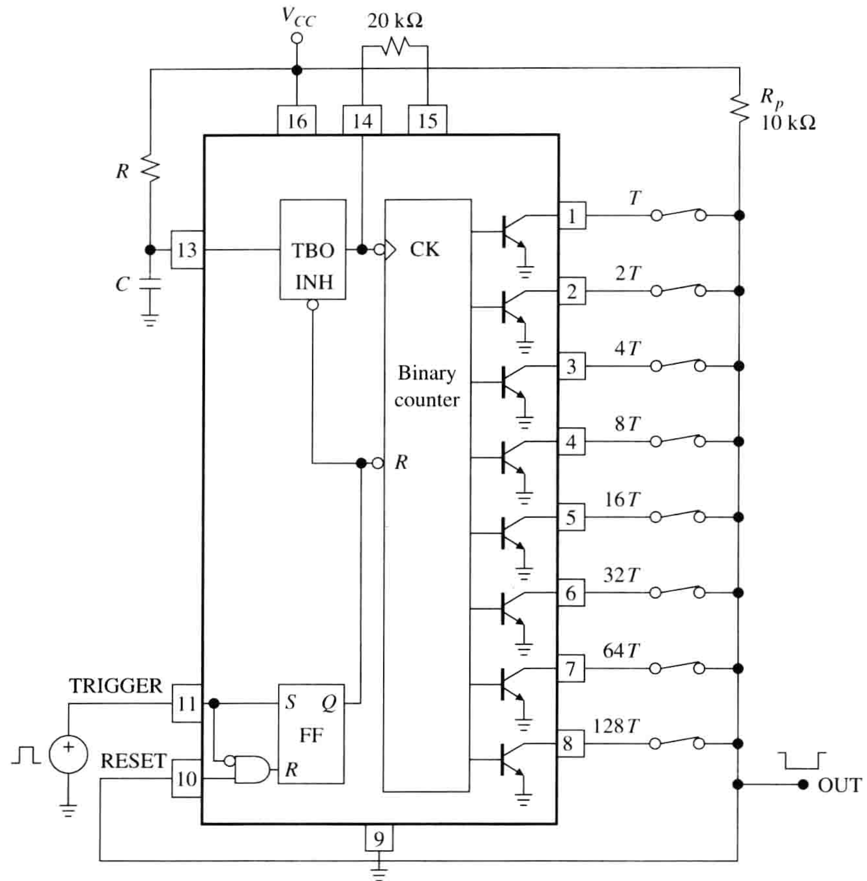
EXAMPLE 10.4. Assuming $V_{CC} = 5$ V in the multivibrator of Example 10.3, find the range of variation of f_0 and $D(\%)$ if the voltage at the CONTROL input is modulated by ac coupling to it an external sine wave with a peak amplitude of 1 V.

Solution. The range of variation of V_{TH} is $(2/3)5 \pm 1$ V, or between 4.333 V and 2.333 V. Substituting into Eq. (10.13) gives $T_L = 4.96\ \mu$s and $7.78\ \mu\text{s} \le T_H \le 31.0\ \mu$s, so we have $27.8\ \text{kHz} \le f_0 \le 78.5$ kHz, and $61.1\% \le D(\%) \le 86.2\%$.

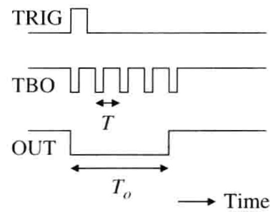
Timer/Counter Circuits

In applications requiring very long delays, the values of the timing components can become impractically large. This drawback is overcome by using components of manageable size and then stretching the multivibrator time scale with a binary counter. This concept is exploited in the popular 2240 timer/counter circuit, as well as other similar devices. As shown in Fig. 10.18, the basic elements of the 2240 are a time-base oscillator (TBO), an 8-bit ripple counter, and a control flip-flop (FF). The TBO is similar to the 555 timer, except that R_B has been eliminated to reduce the external component count, and the comparator thresholds have been changed to $V_{TL} = 0.27V_{CC}$ and $V_{TH} = 0.73V_{CC}$ to make the value of the logarithm in Eq. (10.13) exactly unity. Thus, the time-base is

$$T = RC \tag{10.17}$$



(*a*)



(*b*)

FIGURE 10.18
(*a*) Programmable delay generator using the XR2240 timer/counter. (*b*) Timing diagram. (Courtesy of Exar.)

The binary counter consists of eight toggle flip-flops that are buffered by open-collector BJTs. The desired amount of time stretching is programmed by connecting a suitable combination of counter outputs to a common pullup resistor R_p in a wired-OR configuration. Once a particular combination is selected, the output will be low as long as any one of the selected outputs is low. For instance, connecting only pin 5 to the pullup resistor gives $T_o = 16T$, while connecting pins 1, 3, and 7 gives $T_o = (1+4+64)T = 69T$, where T_o is the duration of the output timing cycle. By suitable choice of the connection pattern, one can program T_o anywhere over the range $T \leq T_o \leq 255T$.

The purpose of the control flip-flop is to translate the external TRIGGER and RESET commands to the proper controls for the TBO and the counter. At power turn-on the circuit comes up in the reset state, where the TBO is inhibited and all open-collector outputs are high. On receiving an external trigger pulse, the control flip-flop goes high and initiates a timing cycle by enabling the TBO and forcing the common output node of the counter low. The TBO will now run until the count programmed by the wired-OR pattern is reached. At this point the output goes high, resetting the control flip-flop and stopping the TBO. The circuit is now in the reset state, awaiting the arrival of the next trigger pulse.

Cascading the counter stages of two or more 2240s makes it possible to achieve truly long delays. For instance, cascading two 8-bit counters yields an effective counter length of 16 bits, which allows T_o to be programmed anywhere in the range from T to over $65 \times 10^3 T$. In this manner, delays of hours, days, or months can be generated using relatively small timing component values. Since the counters do not affect the timing accuracy, the accuracy of T_o depends only on that of T, which is typically around 0.5%. T can be fine-tuned by adjusting R.

10.4 TRIANGULAR WAVE GENERATORS

Triangular waves are generated by alternately charging and discharging a capacitor with a constant current. In the circuit of Fig. 10.19*a* the current drive for C is provided by *OA*, a JFET-input op amp functioning as a floating-load *V-I* converter. The converter receives a two-level drive from a 301 op amp comparator configured as a Schmitt trigger. Because of the inversion introduced by *OA*, the Schmitt trigger must be of the noninverting type. Also shown is a diode clamp to stabilize the Schmitt-trigger output levels at $\pm V_{\text{clamp}} = \pm(V_{Z5} + 2V_{D(\text{on})})$. Consequently, the Schmitt-trigger input thresholds are $\pm V_T = \pm(R_1/R_2)V_{\text{clamp}}$.

Circuit behavior is visualized in terms of the waveforms of Fig. 10.19*b*. Assume at power turn-on ($t = 0$) *CMP* swings to $+V_{\text{sat}}$ so that $v_{\text{SQ}} = +V_{\text{clamp}}$. *OA* converts this voltage to a current of value V_{clamp}/R entering C from the left. This causes v_{TR} to ramp downward. As soon as v_{TR} reaches $-V_T$, the Schmitt trigger snaps and v_{SQ} switches from $+V_{\text{clamp}}$ to $-V_{\text{clamp}}$. *OA* converts this new voltage to a capacitance current of the same magnitude but opposite polarity. Consequently, v_{TR} will now ramp upward. As soon as v_{TR} reaches $+V_T$, the Schmitt trigger snaps again, thus repeating the cycle. Figure 10.19*b* shows also the waveform v_1 at the noninverting input of *CMP*. By the superposition principle, this waveform is a linear combination of v_{TR} and v_{SQ}, and it causes the Schmitt trigger to snap whenever it reaches 0 V.

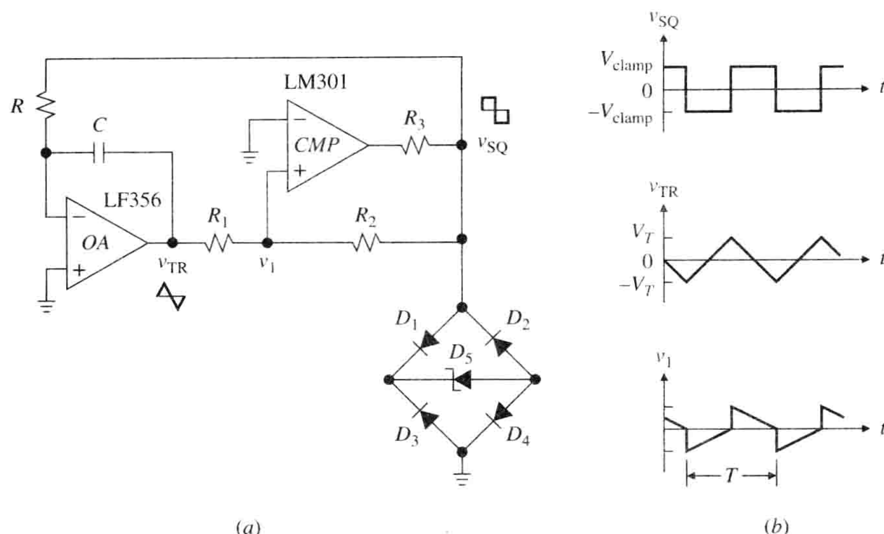


FIGURE 10.19
Basic triangular/square-wave generator.

By symmetry, the time taken by v_{TR} to ramp from $-V_T$ to $+V_T$ is $T/2$. Since the capacitor is operated at constant current, we can apply Eq. (10.2) with $\Delta t = T/2$, $I = V_{clamp}/R$ and $\Delta v = 2V_T = 2(R_1/R_2)V_{clamp}$. Letting $f_0 = 1/T$ gives

$$f_0 = \frac{R_2/R_1}{4RC} \tag{10.18}$$

indicating that f_0 depends only on external components, a desirable feature indeed. As usual, f_0 can be varied continuously by means of R, or in decade steps by means of C. The operating frequency range is limited at the upper end by the SR and GBP of *OA* as well as the speed of response of *CMP*; at the lower end by the size of R and C, as well as the input bias current of *OA* and capacitor leakage. A FET-input op amp is usually a good choice for *OA*, while *CMP* should be an uncompensated op amp or, better yet, a high-speed voltage comparator.

EXAMPLE 10.5. In the circuit of Fig. 10.19*a* specify suitable components for a square wave with peak values of ± 5 V, a triangular wave with peak values of ± 10 V, and f_0 continuously variable from 10 Hz to 10 kHz.

Solution. We need $V_{Z5} = V_{clamp} - 2V_{D(on)} = 5 - 2 \times 0.7 = 3.6$ V, and $R_2/R_1 = V_{clamp}/V_T = 5/10 = 0.5$ (use $R_1 = 20\,\text{k}\Omega$, $R_2 = 10\,\text{k}\Omega$). Since f_0 must be variable over a 1000 : 1 range, implement R with a pot and a series resistance R_s such that $R_{pot} + R_s = 1000R_s$, or $R_{pot} \cong 10^3 R_s$. Use $R_{pot} = 2.5\,\text{M}\Omega$ and $R_s = 2.5\,\text{k}\Omega$. For $R = R_{min} = R_s$ we want $f_0 = f_{0(max)} = 10$ kHz. By Eq. (10.18), $C = 0.5/(10^4 \times 4 \times 2.5 \times 10^3) = 5$ nF. The function of R_3 is to provide current to R, R_2, the diode bridge, and the output load under all operating conditions. Now, $I_{R(max)} = V_{clamp}/R_{min} = 5/2.5 = 2$ mA, and $I_{R_2(max)} = V_{clamp}/R_2 = 5/10 = 0.5$ mA. Imposing a bridge current of 1 mA and allowing for a maximum load current of 1 mA yields $I_{R_3(max)} = 2 + 0.5 + 1 + 1 = 4.5$ mA. Then, $R_3 = (13 - 5)/4.5 = 1.77$ kΩ (use 1.5 kΩ to be safe). For the diode bridge, use a CA3039 diode array.

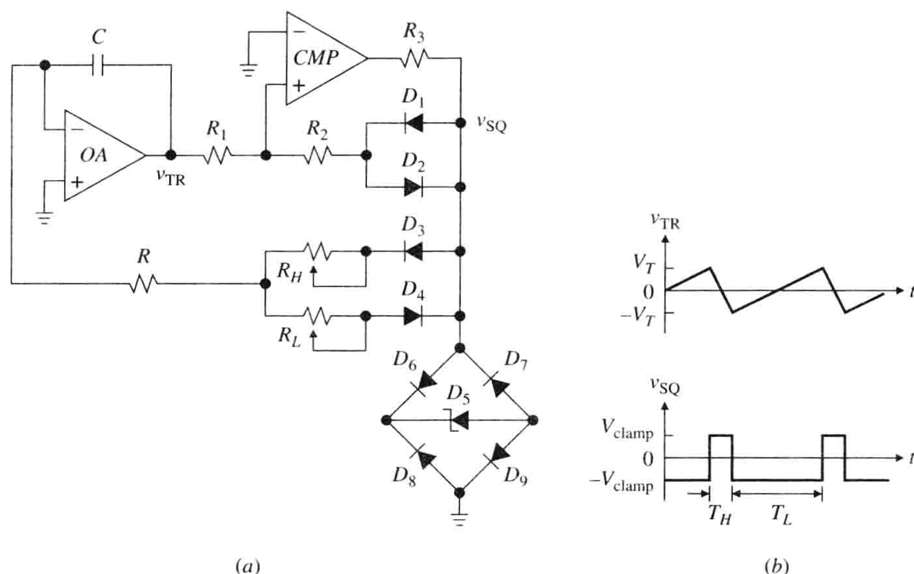


FIGURE 10.20
Triangular wave generator with independently adjustable slopes.

Slope Control

With the modification of Fig. 10.20*a*, the charge and discharge times can be adjusted independently to generate asymmetric waves. With $v_{SQ} = +V_{clamp}$, D_3 is on and D_4 is off, so the discharge current is $I_H = [V_{clamp} - V_{D(on)}]/(R_H + R)$. With $v_{SQ} = -V_{clamp}$, D_3 is off and D_4 is on, and the charge current is $I_L = [V_{clamp} - V_{D(on)}]/(R_L + R)$. The charge and discharge times are found as $C \times 2V_T = I_L T_L$ and $C \times 2V_T = I_H T_H$, respectively. The function of D_1 and D_2 is to compensate for the $V_{D(on)}$ term due to D_3 and D_4. With D_1 and D_2 in place we now have $V_T/R_1 = [V_{clamp} - V_{D(on)}]/R_2$. Combining all the above information yields

$$T_L = 2\frac{R_1}{R_2}C(R_L + R) \qquad T_H = 2\frac{R_1}{R_2}C(R_H + R) \tag{10.19}$$

The frequency of oscillation is $f_0 = 1/(T_H + T_L)$. Note that if one of the slopes is made much steeper than the other, v_{TR} will approach a sawtooth and v_{SQ} a train of narrow pulses.

Voltage-Controlled Oscillator

Many applications require that f_0 be programmable automatically, for instance, via a control voltage v_I. The required circuit, known as a *voltage-controlled oscillator* (VCO), is designed to give $f_0 = kv_I$, $v_I > 0$, where *k* is the sensitivity of the VCO, in hertz per volt.

Figure 10.21 shows a popular VCO realization. Here *OA* is a *V-I* converter that forces *C* to conduct a current linearly proportional to v_I. To ensure capacitor charging

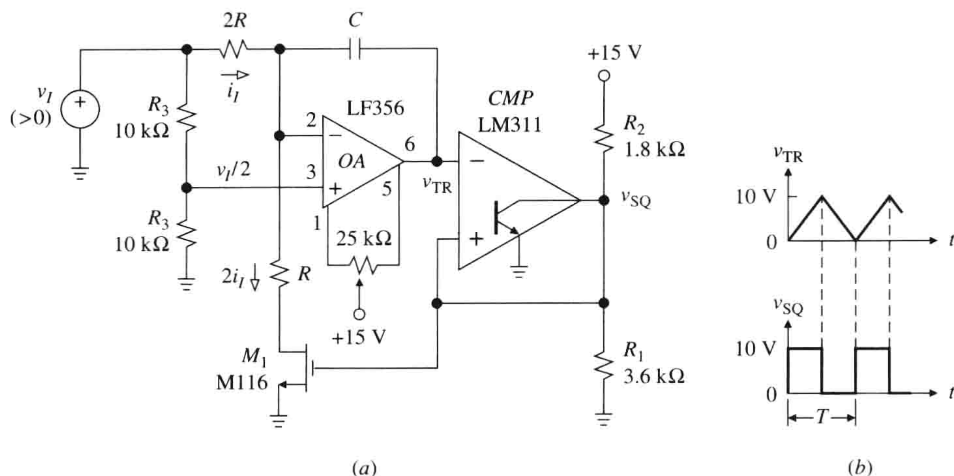


FIGURE 10.21
Voltage-controlled triangular/square-wave oscillator. (Power supplies are ±15 V.)

as well as discharging, this current must alternate between opposite polarities. As we shall see shortly, polarity is controlled via the n-MOSFET switch. Moreover, *CMP* forms a Schmitt trigger whose output levels are $V_{OL} = V_{CE(\text{sat})} \cong 0$ V when the output BJT is saturated, and $V_{OH} = V_{CC}/(1 + R_2/R_1) = 10$ V when the BJT is off. Since the noninverting input is obtained directly from the output, the trigger thresholds are likewise $V_{TL} = 0$ V and $V_{TH} = 10$ V. The circuit operates as follows.

By op amp and voltage-divider action, the voltage at both inputs of *OA* is $v_I/2$, so the current through the $2R$ resistance is at all times $i_I = (v_I - v_I/2)/2R = v_I/4R$. Assume the Schmitt trigger starts in the low state, or $v_{SQ} \cong 0$ V. With a low gate voltage, M_1 is off, so all the current supplied by the $2R$ resistance flows into C, causing v_{TR} to ramp downward.

As soon as v_{TR} reaches $V_{TL} = 0$ V, the Schmitt trigger snaps, causing v_{SQ} to jump to 10 V. With a high gate voltage, M_1 turns on and shorts R to ground, sinking the current $(v_I/2)/R = 2i_I$. Since only half of this current is supplied by the $2R$ resistance, the other half must come from C. Thus, the effect of turning on M_1 is to reverse the current through C without affecting its magnitude. Consequently, v_{TR} is now ramping upward.

As soon as v_{TR} reaches $V_{TH} = 10$ V, the Schmitt trigger snaps back to 0 V, turning off M_1 and reestablishing the conditions of the previous half-cycle. The circuit is therefore oscillating. Using Eq. (10.2) with $\Delta t = T/2$, $I = v_I/(4R)$, and $\Delta v = V_{TH} - V_{TL}$, and then solving for $f_0 = 1/T$ gives

$$f_0 = \frac{v_I}{8RC(V_{TH} - V_{TL})} \tag{10.20}$$

With $V_{TH} - V_{TL} = 10$ V we get $f_0 = kv_I$, $k = 1/(80RC)$. Using, for example, $R = 10$ kΩ, $2R = 20$ kΩ, and $C = 1.25$ nF gives a sensitivity $k = 1$ kHz/V. Then, varying v_I over the range of 10 mV to 10 V sweeps f_0 over the range of 10 Hz to 10 kHz.

The accuracy of Eq. (10.20) is limited at high frequencies by the dynamics of *OA*, *CMP*, and M_1, and at low frequencies by the input bias current and offset voltage of *OA*. To null the latter, set v_I to a low value, say, 10 mV, and then adjust the offset-nulling pot for a 50% duty cycle. Another source of error is the channel resistance $r_{ds(\text{on})}$ of the FET switch. The data sheets of the M116 FET give $r_{ds(\text{on})} = 100\ \Omega$ typical. With $R = 10\ \text{k}\Omega$, this represents an error of only 1%; if desired, this can be eliminated by reducing R from 10 kΩ to 10 kΩ − 100 Ω = 9.9 kΩ.

Triangular-to-Sine Wave Conversion

If a triangular wave is passed through a circuit exhibiting a sinusoidal VTC, as shown in Fig. 10.22*a*, the result is a sine wave. Since nonlinear wave shaping is independent of frequency, this form of sine wave generation is particularly convenient when used in connection with triangular-output VCOs, since the latter offer much wider tuning ranges than Wien-bridge oscillators. Practical wave shapers approximate a sinusoidal VTC by exploiting the nonlinear characteristics of diodes or transistors.[4]

In the circuit of Fig. 10.22*b* a sinusoidal VTC is approximated by suitably over-driving an emitter-degenerated differential pair. Near the zero-crossings of the input, the gain of the pair is approximately linear; however, as either peak is approached, one of the BJTs is driven to the verge of cutoff, where the VTC becomes logarithmic and produces a gradual rounding of the triangular wave. The THD of the output is minimized[4] at about 0.2% for $RI \cong 2.5V_T$ and $V_{im} \cong 6.6V_T$, where V_{im} is the peak amplitude of the triangular wave and V_T is the thermal voltage ($V_T \cong 26$ mV at room temperature). This translates to $RI \cong 65$ mV and $V_{im} \cong 172$ mV, indicating that the triangular wave must be properly scaled to fit the requirements of the wave shaper.

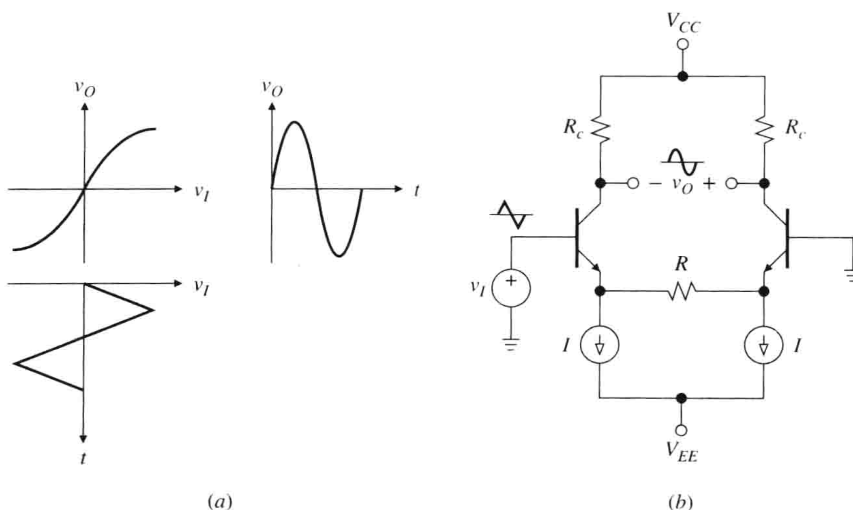


FIGURE 10.22
(*a*) VTC of a triangular-to-sine wave converter. (*b*) Logarithmic wave shaper.

EXAMPLE 10.6. Design a circuit to convert a 1-kHz triangular wave of ±5-V peak values to a ground-referenced sine wave, also of ±5-V peak values, and visualize the input and output waveforms via PSpice. For a change, use the LF411 JFET-input op amp, whose macromodel is available in PSpice's library. Also, assume ±9-V power supplies.

Solution. First, use a voltage divider (R_1 and R_2 in Fig. 10.23*a*) to scale the ±5-V input v_T to a ±172-mV drive for the base of Q_1. Next, use $R_7 = 65\ \Omega$ to implement the condition $(65\ \Omega) \times (1\ \text{mA}) = 2.5V_T = 65$ mV. Finally, use an LF411-based difference amplifier to convert the voltage difference between the collectors to a single-ended output v_S. Judging by the waveforms shown, v_S is a fairly good sine wave representation.

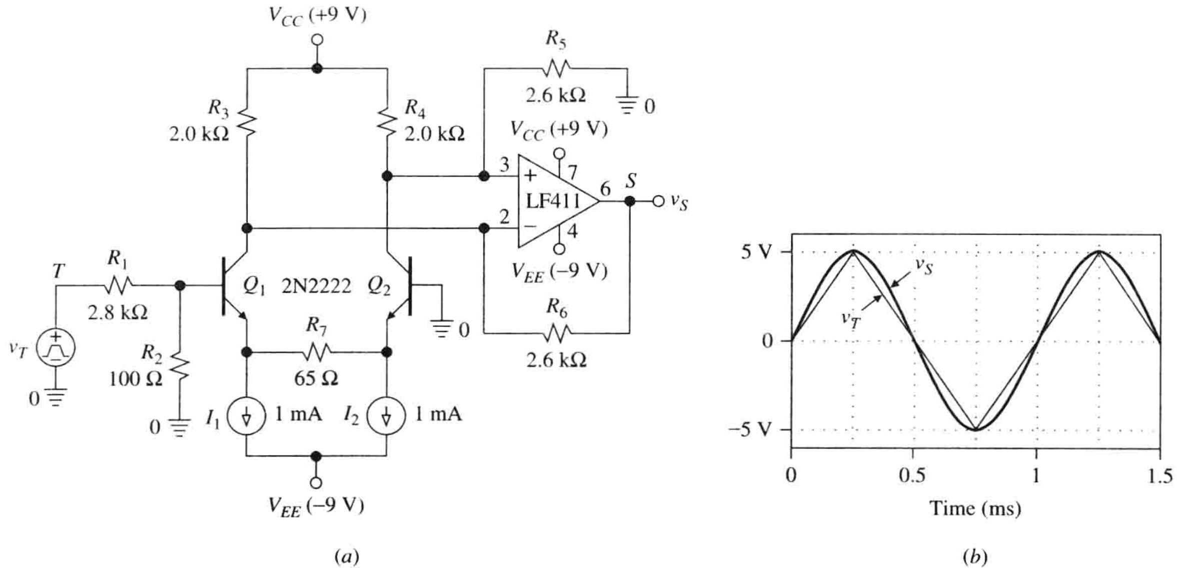


FIGURE 10.23
(*a*) PSpice triangular-to-sine converter and (*b*) input/output waveforms.

10.5 SAWTOOTH WAVE GENERATORS

A sawtooth cycle is generated by charging a capacitor at a constant rate and then rapidly discharging it with a switch. Figure 10.24 shows a circuit utilizing this principle. The current drive for C is provided by *OA*, a floating-load *V-I* converter. In order for v_{ST} to be a positive ramp, i_I must always flow out of the summing junction, or $v_I < 0$. R_2 and R_3 establish the threshold $V_T = V_{CC}/(1 + R_2/R_3) = 5$ V.

At power turn-on ($t = 0$), when C is still discharged, the 311 comparator inputs are $v_P = 0$ V and $v_N = 5$ V, indicating that the output BJT is in saturation and $v_{\text{PULSE}} \cong -15$ V. With a gate voltage this low, the *n*-JFET J_1 is in cutoff, allowing C to charge. As soon as the ensuing ramp v_{ST} reaches V_T, the comparator output BJT goes off, allowing the 2-kΩ resistor to pull v_{PULSE} to ground. This change of state takes place in a snapping fashion because of the positive-feedback action provided by C_1. Since now $v_{GS} = 0$ V, the JFET switch closes and rapidly discharges C, bringing v_{ST} to 0 V.

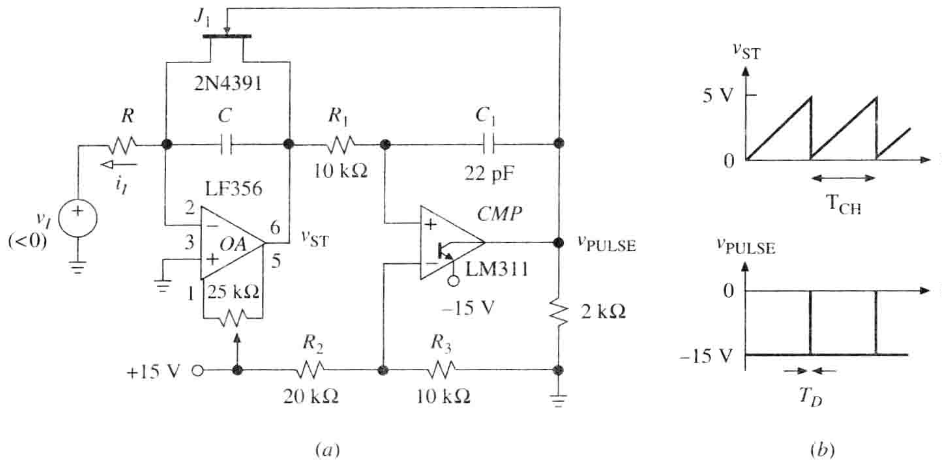


FIGURE 10.24
Voltage-controlled sawtooth/pulse-wave oscillator.

The comparator is prevented from responding immediately to this change in v_{ST} because of the charge accumulated in C_1 during the transition of v_{PULSE} from -15 V to 0 V. This one-shot action, whose duration T_D is proportional to R_1C_1, is designed to ensure that C undergoes complete discharge. With the component values shown, $T_D < 1\ \mu$s. After timing out, v_{PULSE} returns to -15 V, turning J_1 off again and allowing C to resume charging. The cycle, therefore, repeats itself.

The charging time T_{CH} is found using Eq. (10.2) with $\Delta t = T_{CH}$, $I = |v_I|/R$, and $\Delta v = V_T$. Letting $f_0 = 1/(T_{CH} + T_D)$, we obtain

$$f_0 = \frac{1}{RCV_T/|v_I| + T_D} \tag{10.21}$$

As long as $T_D \ll T_{CH}$, this simplifies to

$$f_0 = \frac{|v_I|}{RCV_T} \tag{10.22}$$

indicating that f_0 is linearly proportional to the control voltage v_I. With $R = 90.9\ \text{k}\Omega$ and $C = 2.2$ nF, $f_0 = k|v_I|, k = 1$ kHz/V, so varying v_I from -10 mV to -10 V will sweep f_0 from 10 Hz to 10 kHz. The circuit can also function as a *current-controlled oscillator* (CCO) if we drive it directly with a current sink i_I. Then, $f_0 = i_I/(CV_T)$. A common application of sawtooth CCOs is found in electronic music, where the control current is provided by an exponential *V-I* converter designed for a sensitivity of 1 octave per volt over a 10-decade frequency range, typically from 16.3516 Hz to 16.744 kHz.

Practical Considerations

A good choice for *OA* is a FET-input op amp combining low input bias current, which is critical at the low end of the control range, with good slew-rate performance, which is critical at the high end. The input offset voltage is not critical in the CCO mode;

however, offset nulling may be necessary in the VCO mode. Also, J_1 should exhibit low leakage and low $r_{ds(\mathrm{on})}$.

The high-frequency accuracy of the oscillator is limited by the presence of T_D in Eq. (10.21). The ensuing error can be compensated for by speeding up the capacitor charging time to make up for the delay T_D. This can be achieved by making V_T decrease with frequency, for instance, by coupling v_I, which is negative, to the junction of R_2 and R_3 via a suitable series resistance R_4. It can be proved (see Problem 10.31) that choosing $R_4 = (R_2 \parallel R_3) \times (RC/T_D - 1)$ makes f_0 linearly proportional to $|v_I|$, though at the price of a slight reduction of the sawtooth amplitude at high frequencies.

10.6 MONOLITHIC WAVEFORM GENERATORS

Also called *function generators,* these circuits are designed to provide the basic waveforms with a minimum of external components. The heart of a waveform generator is a VCO that generates the triangular and square waves. Passing the triangular wave through an on-chip wave shaper yields the sine wave, whereas configuring the oscillator for a highly asymmetric duty cycle gives the sawtooth and pulse-train waves. The two most frequent VCO configurations are the *grounded capacitor* and the *emitter-coupled types,*[4] both of which are available either as stand-alone units or as part of complex systems, such as phase-locked loops (PLLs), tone decoders, *V-F* converters, and PWM controllers.

Grounded-Capacitor VCOs

These circuits are based on the principle of charging and discharging a grounded capacitor at rates controlled by programmable current generators. With reference to Fig. 10.25*a*, we note that when the switch *SW* is in the up position, *C* charges at a rate set by the current source i_H. Once v_{TR} reaches the upper threshold V_{TH}, the Schmitt trigger changes state and flips *SW* to the down position, causing *C* to discharge at

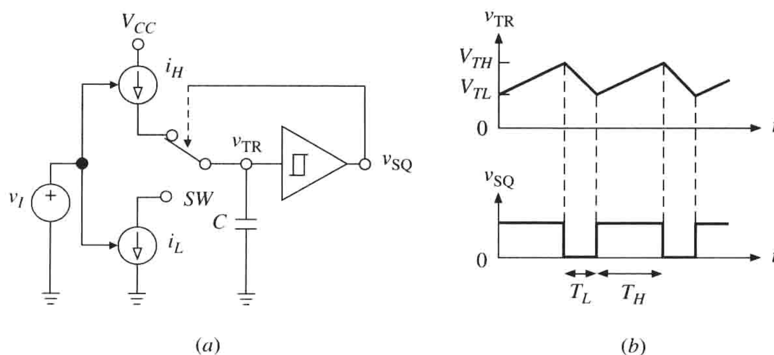


FIGURE 10.25
Grounded-capacitor VCO.

a rate set by the current sink i_L. Once v_{TR} reaches V_{TL}, the trigger changes state again, flipping SW to the up position and repeating the cycle.

To allow for automatic frequency control, i_H and i_L are made programmable via an external control voltage v_I. If the magnitudes of i_L and i_H are equal, the output waveform will be symmetric. Conversely, if one of the currents is made much larger than the other, v_{TR} will approach a sawtooth.

The grounded-capacitor configuration is used in the design of temperature-stable VCOs with operating frequencies up to tens of megahertz. Popular products utilizing this configuration are the NE566 function generator and the ICL8038 precision waveform generator.

The ICL8038 Waveform Generator

In the circuit[5] of Fig. 10.26, Q_1 and Q_2 form two programmable current sources whose magnitudes are set by the external resistors R_A and R_B. The drive for Q_1 and Q_2 is provided by the emitter follower Q_3, which also compensates for their base-emitter voltage drops to yield $i_A = v_I/R_A$ and $i_B = v_I/R_B$, with v_I being referenced to V_{CC} as shown. While i_A is fed to C directly, i_B is diverted to the current mirror Q_4-Q_5-Q_6 where it undergoes polarity reversal as well as amplification by 2 due to the combined action of Q_5 and Q_6. The result is a current sink of magnitude $2i_B$.

The Schmitt trigger is similar to that of the 555 timer, with $V_{TL} = (1/3)V_{CC}$ and $V_{TH} = (2/3)V_{CC}$. When the flip-flop output Q is high, Q_7 saturates and pulls the

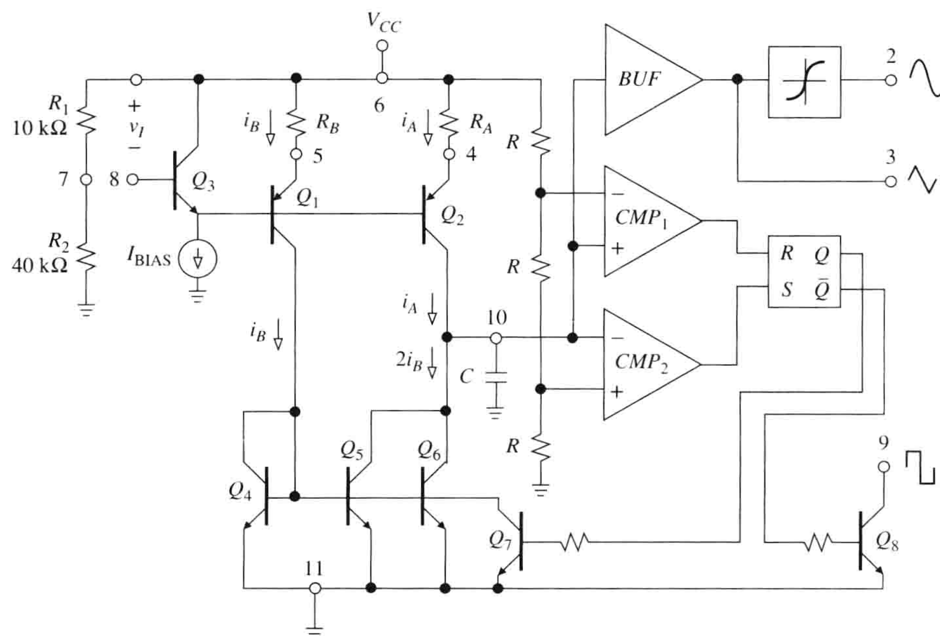


FIGURE 10.26
Simplified circuit diagram of the ICL8038 waveform generator. (Courtesy of Harris Semiconductor.)

bases of Q_5 and Q_6 low, shutting off the current sink. Consequently, C charges at a rate set by $i_H = i_A$. Once the capacitance voltage reaches V_{TH}, CMP_1 fires and clears the flip-flop, turning Q_7 off and enabling the current mirror. The net current out of C is now $i_L = 2i_B - i_A$; as long as $2i_B > i_A$, this current will cause C to discharge. Once V_{TL} is reached, CMP_2 fires and sets the flip-flop, thus repeating the cycle. It can be shown (see Problem 10.36) that

$$f_0 = 3\left(1 - \frac{R_B}{2R_A}\right)\frac{v_I}{R_A C V_{CC}} \qquad D(\%) = 100\left(1 - \frac{R_B}{2R_A}\right) \tag{10.23}$$

With $R_A = R_B = R$ the circuit yields symmetric waveforms with $f_0 = kv_I$, $k = 1.5/RCV_{CC}$. As shown in the figure, the device is also equipped with a unity-gain buffer to isolate the waveform developed across C, a wave shaper to convert the triangular wave to a low-distortion sine wave, and an open-collector transistor (Q_8) to provide, with the help of an external pullup resistor, a square-wave output.

Figure 10.27 shows the wave shaper[5] utilized in the 8038. The circuit is known as a *breakpoint wave shaper* because it uses a set of breakpoints at designated signal levels to fit a nonlinear VTC by a piecewise linear approximation. The circuit, designed to process triangular waves alternating between $(1/3)V_{CC}$ and $(2/3)V_{CC}$, uses the resistive strings shown at the right to establish two sets of breakpoint voltages symmetric about the midrange value of $(1/2)V_{CC}$. These voltages are then buffered by the even-numbered emitter-follower BJTs. The circuit works as follows.

For v_I near $(1/2)V_{CC}$, all odd-numbered BJTs are off, giving $v_O = v_I$. Consequently, the initial slope of the VTC is $a_0 = \Delta v_O/\Delta v_I = 1$ V/V. As v_I is increased to the first breakpoint, the common-base BJT Q_1 goes on and loads down the source, changing the VTC slope from a_0 to $a_1 = 10/(1 + 10) = 0.909$ V/V. Further increasing v_I to the second breakpoint turns Q_3 on, changing the slope to $a_2 = (10 \parallel 2.7)/[1 + (10 \parallel 2.7)] = 0.680$ V/V. The process is repeated for the remaining breakpoints above $(1/2)V_{CC}$ as well for the corresponding breakpoints below $(1/2)V_{CC}$. By progressively reducing the slope as v_I moves away from its midrange value, the circuit approximates a sinusoidal VTC with THD levels around 1% or less. We observe that the even- and odd-numbered BJTs associated with each breakpoint are complementary to each other. This results in a first-order cancellation of the corresponding base-emitter voltage drops, yielding more predictable and stable breakpoints.

Basic 8038 Applications[5]

In the basic connection of Fig. 10.28 the control voltage v_I is derived from V_{CC} via the internal voltage divider R_1 and R_2 (see Fig. 10.26), so $v_I = (1/5)V_{CC}$. Inserting into Eq. (10.23) gives

$$f_0 = \frac{0.3}{RC} \qquad D(\%) = 50\% \tag{10.24}$$

indicating that f_0 is independent of V_{CC}, a desirable feature as we know. By proper choice of R and C, the circuit can be made to oscillate at any frequency from 0.001 Hz to 1 MHz. The thermal drift of f_0 is typically 50 ppm/°C. For optimum performance, confine i_A and i_B within the 1-μA to 1-mA range.

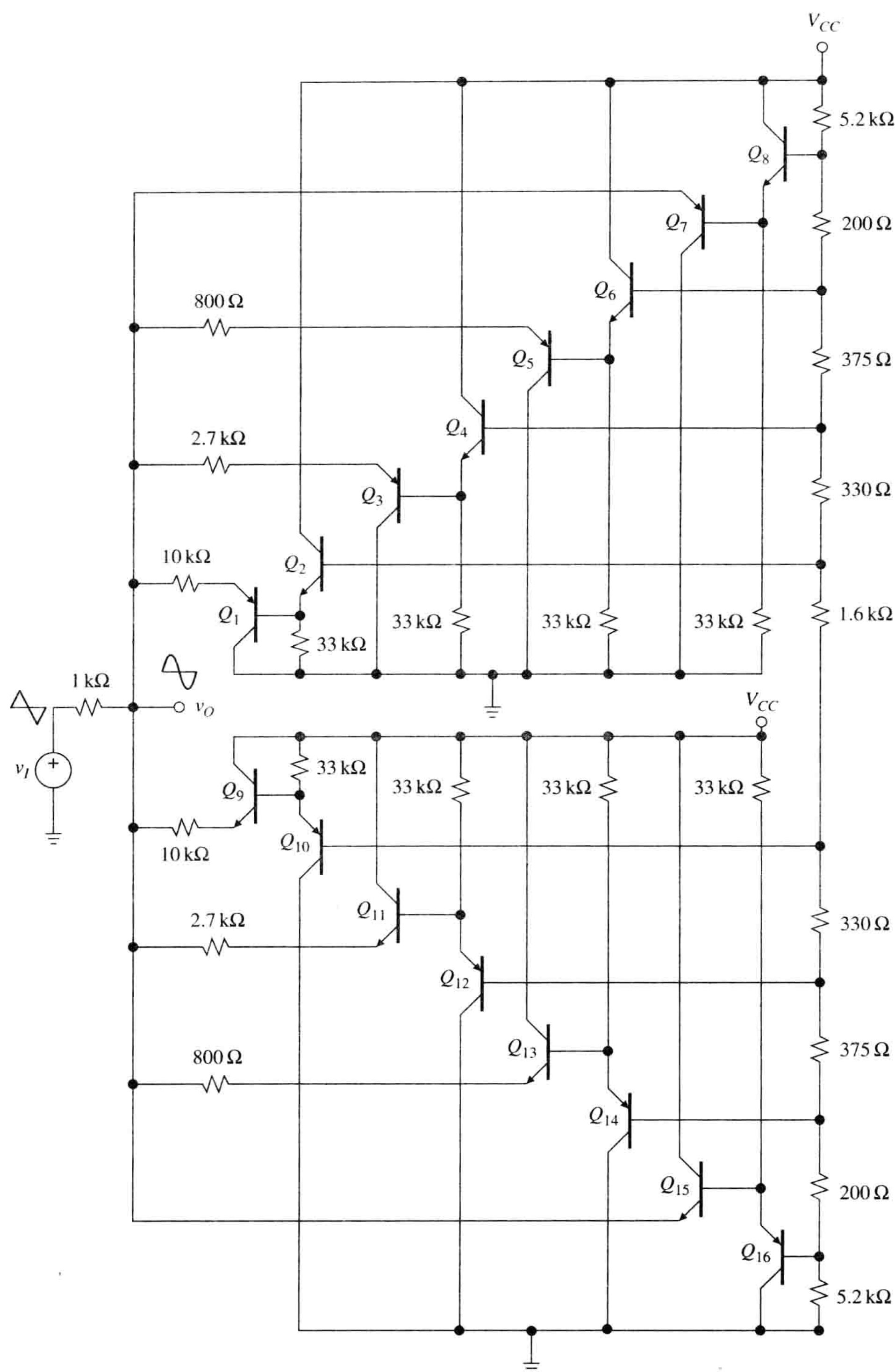


FIGURE 10.27
Breakpoint wave shaper. (Courtesy of Harris Semiconductor.)

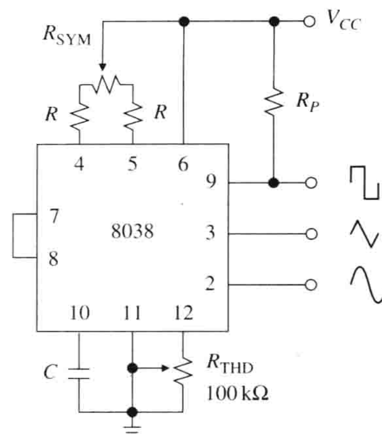


FIGURE 10.28
Basic ICL8038 connection for fixed-frequency, 50% duty cycle operation. (Courtesy of Harris Semiconductor.)

For perfect symmetry it is crucial that i_L and i_H be exactly in a 2:1 ratio. By adjusting R_{SYM}, one can keep the distortion level of the sine wave near 1%. Connecting a 100-kΩ pot between pins 12 and 11 allows one to control the degree of balance of the wave shaper to further reduce the THD.

As mentioned, the square-wave output is of the open-collector type, so a pullup resistor R_p is needed. The peak-to-peak amplitudes of the square, triangular, and sine waves are V_{CC}, $0.33V_{CC}$, and $0.22V_{CC}$, respectively. All three waves are centered at $V_{CC}/2$. Powering the 8038 from split supplies makes the waves symmetric about ground.

EXAMPLE 10.7. Assuming $V_{CC} = 15$ V in the circuit of Fig. 10.28, specify suitable components for $f_0 = 10$ kHz.

Solution. Impose $i_A = i_B = 100\ \mu$A, which is well within the recommended range. Then, $R = (15/5)/0.1 = 30$ kΩ, and $C = 0.3/(10 \times 10^3 \times 30 \times 10^3) = 1$ nF. Use $R_p = 10$ kΩ, and use $R_{SYM} = 5$ kΩ to allow for a ±20% symmetry adjustment. Then, recalculate R as $30 - 5/2 = 27.5$ kΩ (use 27.4 kΩ). To calibrate the circuit, adjust R_{SYM} so that the square wave has $D(\%) = 50\%$, and R_{THD} until the THD of the sine wave is minimized.

Varying the voltage of pin 8 provides automatic frequency sweeps. The fact that the control voltage must be referenced to the V_{CC} rail is annoying in certain applications. This can be avoided by powering the 8038 between ground and a negative supply, as in Fig. 10.29. Also shown in the diagram is an op amp that converts the control voltage v_I to a current i_I, which then splits evenly between Q_1 and Q_2. This scheme also eliminates any errors stemming from imperfect cancellation of the base-emitter voltage drops of Q_3 and the Q_1-Q_2 pair. For accurate *V-I* conversion, the input offset voltage of the op amp must be nulled. The circuit shown is designed to give $i_I = v_I/(5\ \text{k}\Omega)$ over a 1000:1 range, and is calibrated as follows: (*a*) with $v_I = 10.0$ V and the wiper of R_3 set in the middle, adjust R_2 for $D(\%) = 50\%$; hence,

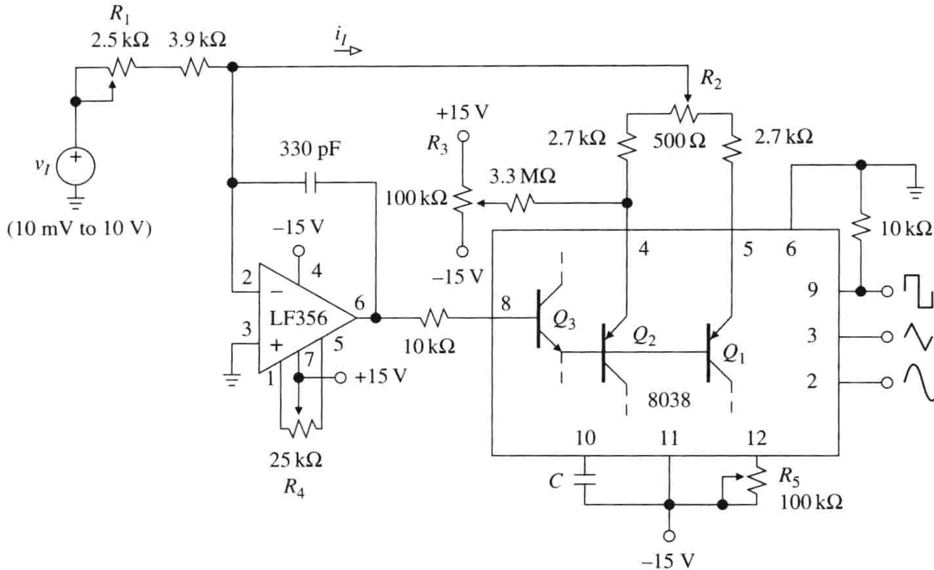


FIGURE 10.29
The ICL8038 as a linear voltage-controlled oscillator. (Courtesy of Harris Semiconductor.)

adjust R_1 for the desired full-scale frequency f_{FS}; (*b*) with $v_I = 10.0$ mV, adjust R_4 for $f_0 = f_{FS}/10^3$; hence, adjust R_3 for $D(\%) = 50\%$; repeat the adjustment of R_4, if necessary; (*c*) with $v_I = 1$ V, adjust R_5 for minimum THD.

Emitter-Coupled VCOs

These VCOs use a pair of cross-coupled Darlington stages and an emitter-coupling timing capacitor, as shown[4] in Fig. 10.30*a*. The two stages are biased with matched emitter currents, and their collector swings are constrained to just one diode voltage drop by clamps D_1 and D_2.

The cross-coupling between the two stages ensures that either Q_1-D_1 or Q_2-D_2 (but not both) are conducting at any given time. This bistable behavior is similar to cross-coupled inverters in flip-flop realizations. Unlike flip-flops, however, the capacitive coupling between the emitters causes the circuit to alternate between its two states in astable-multivibrator fashion. During any half cycle, the capacitor plate connected to the stage that is on remains at a constant potential, while the plate connected to the stage that is off ramps downward at a rate set by i_I. As the ramp approaches the emitter conduction threshold of the corresponding BJT, the latter goes on, forcing the other BJT to go off because of the positive-feedback action stemming from cross-coupling. Thus, C is alternately charged and discharged at a rate set by i_I.

Circuit operation is better visualized by tracing through the waveforms of Fig. 10.30*b*. Note that the emitter waveforms are identical except for a half-cycle delay. Feeding them to a high input-impedance difference amplifier yields a symmetric triangular wave with a peak-to-peak amplitude of two base-emitter voltage drops.

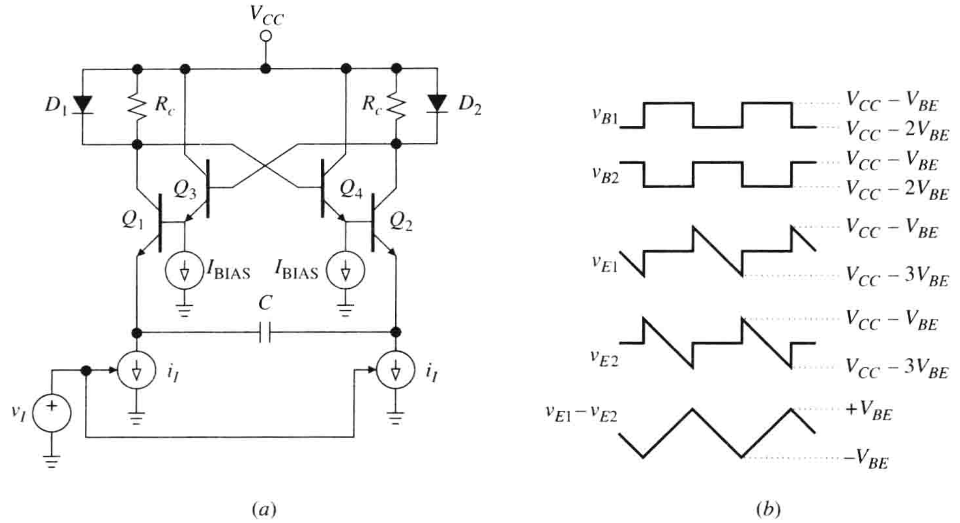


FIGURE 10.30
Emitter-coupled VCO.

The frequency of oscillation is found via Eq. (10.2) with $\Delta t = T/2$ and $\Delta v = 2V_{BE}$. Letting $f_0 = 1/T$ gives

$$f_0 = \frac{i_I}{4CV_{BE}} \tag{10.25}$$

indicating the CCO capability of the circuit.

The emitter-coupled oscillator enjoys a number of advantages: (*a*) it is simple and symmetric, (*b*) it lends itself to automatic frequency control, and (*c*) it is inherently capable of high-frequency operation since it consists of nonsaturating *npn*-BJTs. In its basic form of Fig. 10.30*a*, however, it suffers from a major drawback, namely, the thermal drift of V_{BE}, which is typically -2 mV/°C. There are various methods[4] of stabilizing f_0 with temperature. One method makes i_I proportional to V_{BE} to render their ratio temperature-independent. Popular devices utilizing this technique are the PLLs of the NE560 and XR-210/15 types. Other methods modify the basic circuit to eliminate the V_{BE} term altogether. Though the increased circuit complexity lowers the upper end of the usable frequency range, these methods achieve thermal drifts as low as 20 ppm/°C. Popular products using this approach are the XR2206/07 monolithic function generators and the AD537 *V-F* converter.

The XR2206 Function Generator

This device uses an emitter-coupled CCO to generate the triangular and square waves, and a logarithmic wave shaper to convert the triangle to the sine wave.[4] The CCO parameters are designed so that when the circuit is connected in the basic

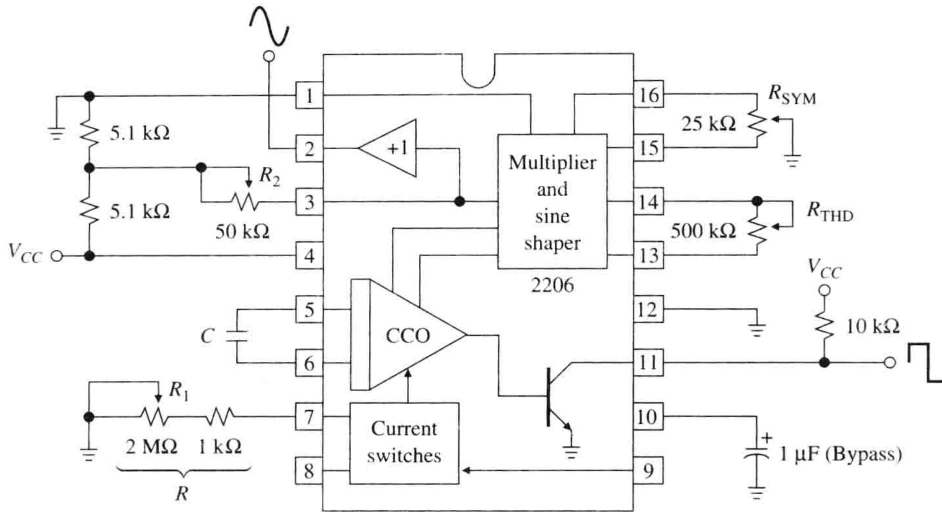


FIGURE 10.31
Basic XR2206 connection for low-distortion sine wave generation. (Courtesy of Exar.)

configuration of Fig. 10.31, the frequency of oscillation is

$$f_0 = \frac{1}{RC} \tag{10.26}$$

The operating frequency range is from 0.01 Hz to more than 1 MHz, with a typical thermal stability of 20 ppm/°C. The recommended range for R is from 1 kΩ to 2 MΩ, and the optimum range is 4 kΩ to 200 kΩ. Varying R with a pot, as shown, allows for a 2000 : 1 sweep of f_0. Symmetry and distortion adjustments are provided, respectively, by R_{SYM} and R_{THD}. With proper calibration the circuit can achieve THD $\cong$ 0.5%.

The amplitude and offset of the sine wave are set by the resistive network external to pin 3. Denoting the equivalent resistance seen by this pin as R_3, the peak amplitude is approximately 60 mV for every kilohm of R_3. For instance, with the wiper of R_2 set in the middle, the peak amplitude of the sine wave is $[25 + (5.1 \parallel 5.1)] \times (60 \text{ mV}) \cong 1.65$ V. The sine wave offset is the same as the dc voltage established by the external network. With the components shown, this is $V_{CC}/2$.

Open circuiting pins 13 and 14 disables the rounding action by the wave shaper so that the output waveform becomes triangular. Its offset is the same as that of the sine wave; however, its peak amplitude is approximately twice as large. The square-wave output is of the open-collector type, hence, a pullup resistor is required.

Figure 10.32 shows another widely used 2206 configuration, which exploits the device's ability to operate with two separate timing resistances R_1 and R_2. With control pin 9 open-circuited or driven high, only R_1 is active and the circuit oscillates at $f_1 = 1/(R_1C)$; similarly, with pin 9 driven low, only R_2 is active and the circuit oscillates at $f_2 = 1/(R_2C)$. Thus, frequency can be keyed between two levels, often referred to as *mark* and *space frequencies,* whose values are set independently by R_1 and R_2. *Frequency shift keying* (FSK) is a widely used method of transmitting

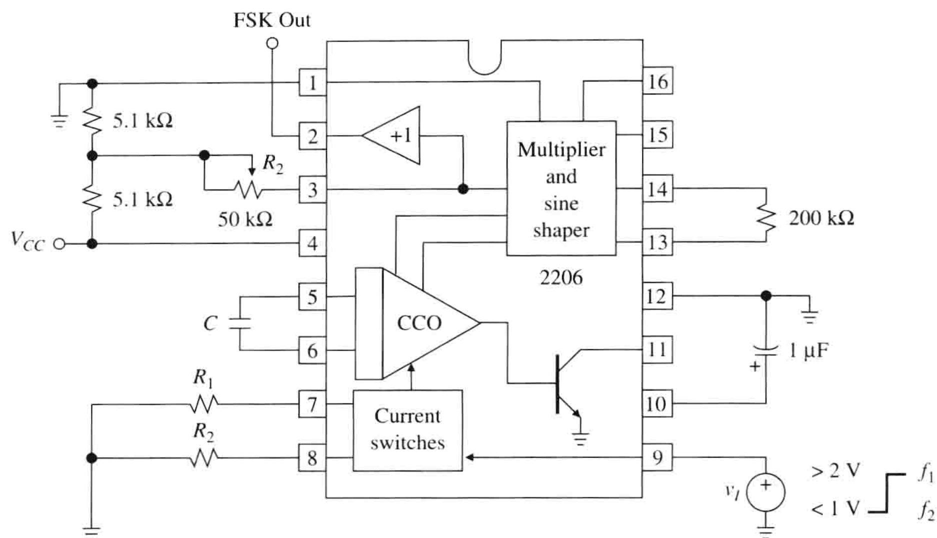


FIGURE 10.32
Sinusoidal FSK generator. (Courtesy of Exar.)

data over telecommunication links. If the FSK control signal is obtained from the square-wave output, R_1 and R_2 will be active on alternate half-cycles of oscillation. This feature can be exploited to configure the 2206 as a sawtooth/pulse generator.

10.7 *V-F* AND *F-V* CONVERTERS

The function of a *voltage-to-frequency converter* (VFC) is to accept an analog input v_I and generate a pulse train with frequency

$$f_O = kv_I \tag{10.27}$$

where k is the VFC *sensitivity*, in hertz per volt. As such, the VFC provides a simple form of analog-to-digital conversion. The primary reason for this type of conversion is that a pulse train can be transmitted and decoded much more accurately than an analog signal, especially if the transmission path is long and noisy. If electrical isolation is also desired, it can be accomplished without loss of accuracy using inexpensive optocouplers or pulse transformers. Moreover, combining a VFC with a binary counter and digital readout provides a low-cost digital voltmeter.[6]

VFCs usually have more stringent performance specifications than VCOs. Typical requirements are (*a*) wide dynamic range (four decades or more), (*b*) the ability to operate to relatively high frequencies (hundreds of kilohertz, or higher), (*c*) low linearity error (less than 0.1% deviation from the straight line going from zero to the full scale), (*d*) high scale-factor accuracy and stability with temperature and supply voltage. The output waveform, on the other hand, is of secondary concern as long as its levels are compatible with standard logic signals. VFCs fall into two categories: *wide-sweep multivibrators* and *charge-balancing VFCs*.[4]

Wide-Sweep Multivibrator VFCs

These circuits are essentially voltage-controlled astable multivibrators designed with VFC performance specifications in mind. The multivibrator is usually a temperature-stabilized version of the basic CCO concept of Fig. 10.30. A popular product[7] in this category is the AD537 shown in Fig. 10.33. The op amp and Q_1 form a buffer *V-I* converter that converts v_I to the current drive i_I for the CCO according to $i_I = v_I/R$. The CCO parameters have been chosen so that $f_O = i_I/(10C)$, or

$$f_O = \frac{v_I}{10RC} \tag{10.28}$$

This relationship holds fairly accurately over a dynamic range of at least four decades, up to a full-scale current of 1 mA and a full-scale frequency of 100 kHz. For instance, with $C = 1$ nF, $R = 10$ kΩ, and $V_{CC} = 15$ V, varying v_I from 1 mV to 10 V varies i_I from 0.1 μA to 1 mA and f_O from 10 Hz to 100 kHz. To minimize the *V-I* conversion error at the low end of the range, the op amp input offset error is nulled internally via R_{OS}. With a capacitor of suitable quality (polystyrene or NPO ceramic for low thermal drift and low dielectric absorption), the linearity error ratings are 0.1% typical for $f_O \leq 10$ kHz, 0.15% typical for $f_O \leq 100$ kHz.

Though the figure shows the connection for $v_I > 0$, we can easily configure the device for $v_I < 0$ by grounding the noninverting input of the op amp, lifting the left terminal of *R* off ground, and applying v_I there. The device can also function as a current-to-frequency converter (CFC) if we make the control current flow out of the inverting input node. For instance, grounding pin 5 and replacing *R* by a photodetector diode current sink will convert light intensity to frequency.

The AD537 also includes an on-chip precision voltage reference to stabilize the CCO scale factor. This yields a typical thermal stability of 30 ppm/°C. To further enhance the versatility of the device, two nodes of the reference circuitry are made available to the user, namely, V_R and V_T. Voltage V_R is a stable 1.00-V voltage reference. Obtaining v_I from pin 7 in Fig. 10.33 yields $f_O = 1/(10RC)$, and if *R* is



FIGURE 10.33
The AD537 voltage-to-frequency converter. (Courtesy of Analog Devices.)

a resistive transducer, such as a photoresistor or a thermistor, it will convert light or temperature to frequency.

Voltage V_T is a voltage linearly proportional to absolute temperature T as $V_T = (1 \text{ mV/K})T$. For instance, at $T = 25\ ^\circ\text{C} = 298.2$ K we have $V_T = 298.2$ mV. If v_I is derived from pin 6 in Fig. 10.33, then $f_O = T/(RC \times 10^4\text{K})$, indicating that the circuit converts absolute temperature to frequency. For instance, with $R = 10\ \text{k}\Omega$ and $C = 1$ nF, the sensitivity is 10 Hz/K. Other temperature scales, such as Celsius and Fahrenheit, can be accommodated by suitably offsetting the input range with the help of V_R.

EXAMPLE 10.8. In the circuit of Fig. 10.34 specify suitable components to yield Celsius-to-frequency conversion with a sensitivity of 10 Hz/°C; then outline the calibration procedure.

Solution. For $T = 0\ ^\circ\text{C} = 273.2$ K we have $V_T = 0.2732$ V and we want $f_O = 0$. Thus, R_3 must develop a 0.2732-V drop. Imposing $0.2732/R_3 = (1.00 - 0.2732)/R_2$ yields $R_2 = 2.66R_3$. For a sensitivity of 10 Hz/°C we want $10 = 1/10^4RC$, where $R = R_1 + (R_2 \parallel R_3)$ is the effective resistance seen by Q_1. Let $C = 3.9$ nF; then $R = 2.564\ \text{k}\Omega$. Let $R_3 = 2.74\ \text{k}\Omega$; then $R_2 = 2.66 \times 2.74 = 7.29\ \text{k}\Omega$ (use 6.34 kΩ in series with a 2-kΩ pot). Finally, $R_1 = 2.564 - (2.74 \parallel 7.29) = 572\ \Omega$ (use 324 Ω in series with a 500-Ω pot).

To calibrate, place the IC in a 0 °C environment and adjust R_2 so that the circuit is barely oscillating, say, $f_O \cong 1$ Hz. Then move the IC to a 100 °C environment and adjust R_1 for $f_O = 1.0$ kHz.

Figure 10.34 shows another useful feature of the AD537, namely, the ability to transmit information over a twisted pair. This pair serves the dual purpose of supplying power to the device and carrying frequency data in the form of current

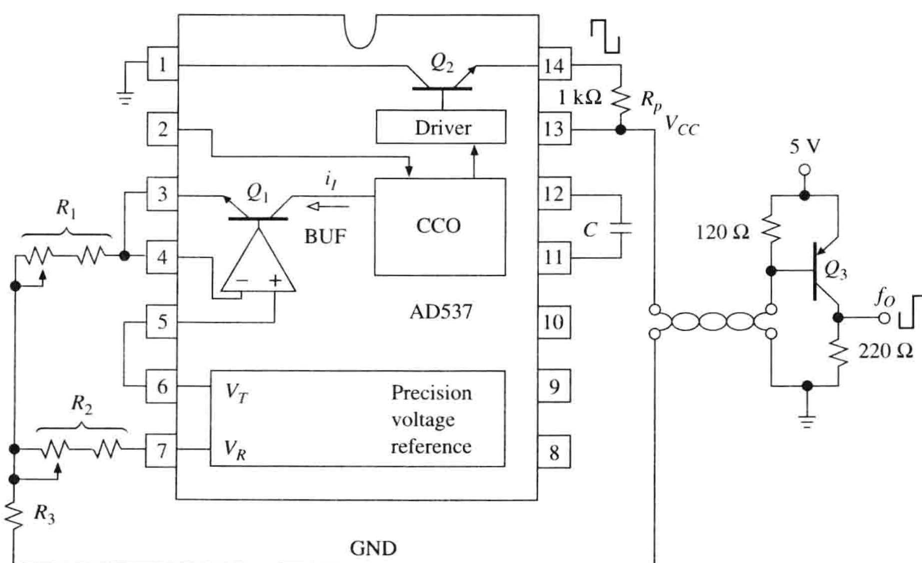


FIGURE 10.34
AD537 application as a temperature-to-frequency converter with two-wire transmission. (Courtesy of Analog Devices.)

modulation. With the parameter values shown, the current drawn by the AD537 alternates between about 1.2 mA during the half-cycle in which Q_2 is off, and $1.2 + [5 - V_{EB3(\text{sat})} - V_{CE2(\text{sat})}]/R_p \cong 1.2 + (5 - 0.8 - 0.1)/1 = 5.3$ mA during the half-cycle in which Q_2 is on. This current difference is sensed by Q_3 as a voltage drop across the 120-Ω resistance. This drop is designed to be low enough to keep Q_3 in cutoff when the current is 1.2 mA, yet large enough to drive Q_3 in saturation when the current is 5.3 mA. Consequently, Q_3 reconstructs a 5-V square wave at the receiving end. The ripple of about 0.5 V appearing across the 120-Ω resistance does not affect the performance of the AD537, thanks to its high PSRR.

Charge-Balancing VFCs

The charge-balancing technique[8] supplies a capacitor with continuous charge at a rate that is linearly proportional to the input voltage v_I, while simultaneously pulling discrete charge packets out of the capacitor at a rate f_O such that the net charge flow is always zero. The result is $f_O = kv_I$. Figure 10.35 illustrates the principle using the VFC32 *V-F* converter.

OA converts v_I to a current $i_I = v_I/R$ flowing into the summing junction; the value of R is chosen such that we always have $i_I < 1$ mA. With *SW* open, i_I flows into C_1 and causes v_1 to ramp downward. As soon as v_1 reaches 0 V, *CMP* fires and triggers a precision one-shot that closes *SW* and turns on Q_1 for a time interval

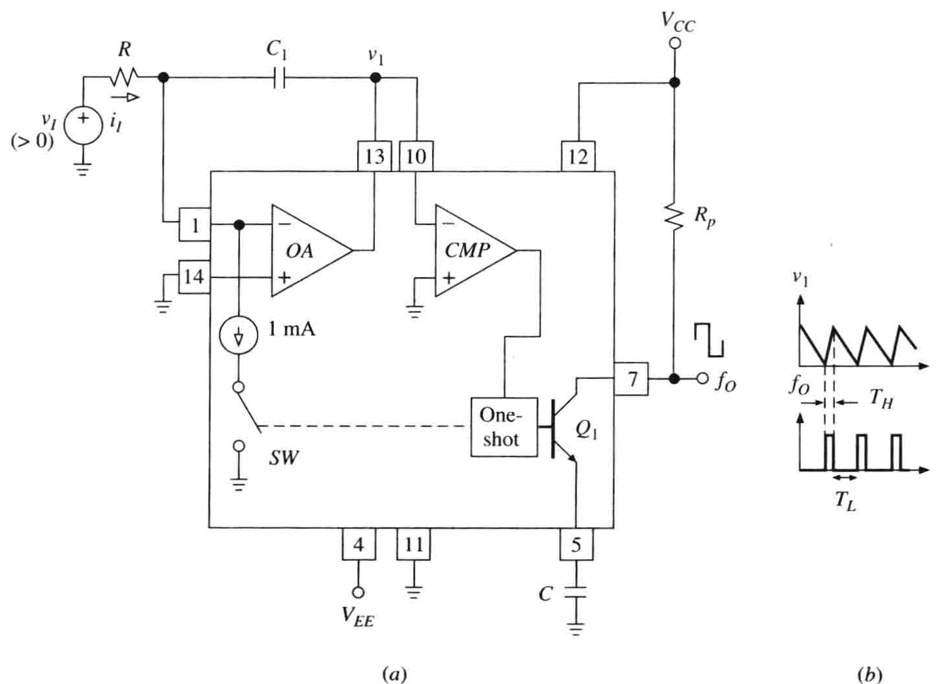


FIGURE 10.35
The VFC32 voltage-to-frequency converter. (Courtesy of Texas Instruments.)

T_H set by C. The one-shot, whose details have been omitted for simplicity, uses a threshold of 7.5 V and a charging current of 1 mA to give

$$T_H = \frac{7.5 \text{ V}}{1 \text{ mA}} C \tag{10.29}$$

The closure of *SW* causes a net current of magnitude $(1 \text{ mA} - i_I)$ to flow out of the summing junction of *OA*. Consequently, during T_H, v_1 ramps upward by an amount $\Delta v_1 = (1 \text{ mA} - i_I)T_H/C_1$. After the one-shot times out, *SW* is opened and v_1 resumes ramping downward at a rate again set by i_I. The time T_L it takes for v_1 to return to zero is such that $T_L = C_1 \Delta v_1 / i_I$. Eliminating Δv_1 and letting $f_O = 1/(T_L + T_H)$ gives, with the help of Eq. (10.29),

$$f_O = \frac{v_I}{7.5RC} \tag{10.30}$$

where f_O is in hertz, v_I in volts, R in ohms, and C in farads. As desired, f_O is linearly proportional to v_I. Moreover, the duty cycle $D(\%) = 100 \times T_H/(T_H + T_L)$ is readily found to be

$$D(\%) = 100 \frac{v_I}{R \times 1 \text{ mA}} \tag{10.31}$$

and it is also proportional to v_I. For best linearity, the data sheets recommend designing for a maximum duty cycle of 25%, which corresponds to $i_{I(\max)} = 0.25$ mA.

The absence of C_1 from the above equations indicates that the tolerance and drift of this capacitor are not critical, so its value can be chosen arbitrarily. However, for optimum performance, the data sheets recommend using the value of C_1 that yields $\Delta v_1 \cong 2.5$ V. C, on the other hand, does appear in Eq. (10.30), so it must be a low-drift type, such as NPO ceramic. If C and R have equal but opposing thermal coefficients, the overall drift can be reduced to as little as 20 ppm/°C. For accurate operation to low values of v_I, the input offset voltage of *OA* must be nulled.

The VFC32 offers a 6-decade dynamic range with typical linearity errors of 0.005%, 0.025%, and 0.05% of full-scale reading for full-scale frequencies of 10 kHz, 100 kHz, and 500 kHz, respectively. Though Fig. 10.35 shows the connection for $v_I > 0$, the circuit is readily configured for $v_I < 0$ or for current-input operation in a manner similar to the AD537 discussed above.

EXAMPLE 10.9. In the circuit of Fig. 10.35 specify components so that a full-scale input of 10 V yields a full-scale output of 100 kHz. The circuit is to have provisions for offset voltage nulling as well as full-scale adjustment.

Solution. We have $T = 1/10^5 = 10$ μs. For $D(\%)_{\max} = 25\%$ use $T_H = 2.5$ μs. By Eq. (10.29), $C = 2.5 \times 10^{-6} \times 10^{-3}/7.5 = 333$ pF (use a 330-pF NPO capacitor with 1% tolerance). By Eq. (10.30), $R = 10/(7.5 \times 330 \times 10^{-12} \times 10^5) = 40.4$ kΩ (use a 34.8-kΩ, 1% metal-film resistor in series with a 10-kΩ cermet pot for full-scale adjustment). Imposing $\Delta v_{1(\max)} = 2.5$ V yields $C_1 = (10^{-3} \times 2.5 \times 10^{-6})/2.5 = 1$ nF.

To null the input offset voltage of *OA*, use the scheme of Fig. 5.20*b* with $R_A = 62$ Ω, $R_B = 150$ kΩ, and $R_C = 100$ kΩ. The calibration is similar to that of Example 10.8.

Frequency-to-Voltage Conversion

The *frequency-to-voltage converter* (FVC) performs the inverse operation, namely, it accepts a periodic waveform of frequency f_I and yields an analog output voltage

$$v_O = kf_I \tag{10.32}$$

where k is the FVC sensitivity, in volts per hertz. FVCs find application as tachometers in motor speed control and rotational measurements. Moreover, they are used in conjunction with VFCs to convert the transmitted pulse train back to an analog voltage.

A charge-balancing VFC can easily be configured as an FVC by applying the periodic input to the comparator and deriving the output from the op amp, which now has the resistance R in the feedback path (see Fig. 10.36). The input signal usually requires proper conditioning to produce a voltage with reliable zero-crossings for *CMP*. Shown in the figure is a high-pass network to accommodate inputs of the TTL and CMOS type. On each negative spike of v_1, *CMP* triggers the one-shot, closing *SW* and pulling 1 mA out of C_1 for a duration T_H as given in Eq. (10.29). In response to this train of current pulses, v_O builds up until the current pulled out of the summing junction of *OA* in 1-mA packets is exactly counterbalanced by that injected by v_O via R continuously, or $f_I \times 10^{-3} \times T_H = v_O/R$. Solving for v_O and using Eq. (10.29) gives

$$v_O = 7.5RCf_I \tag{10.33}$$

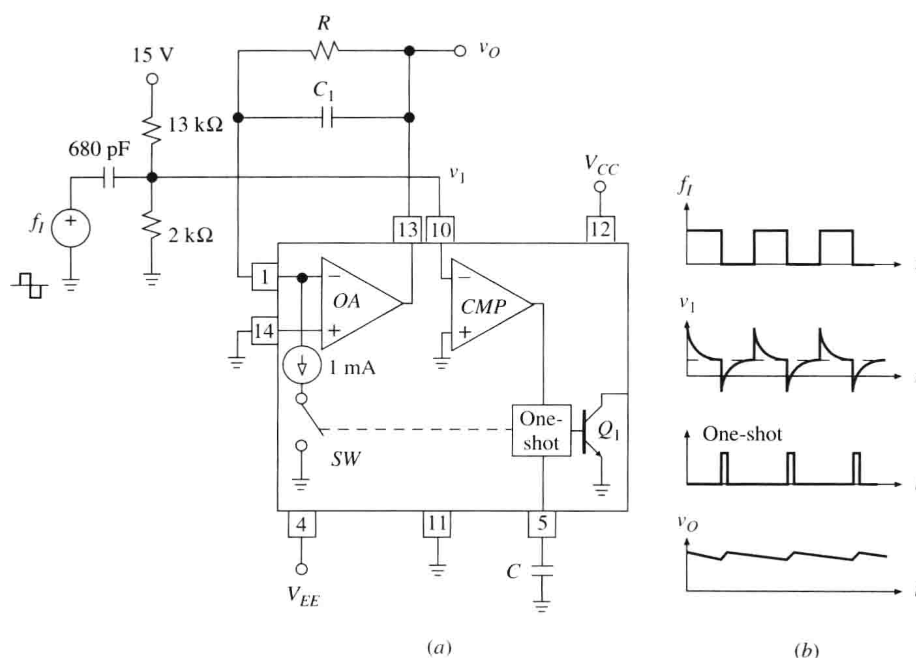


FIGURE 10.36
VFC connection for frequency-to-voltage conversion, and corresponding waveforms. (Courtesy of Texas Instruments.)

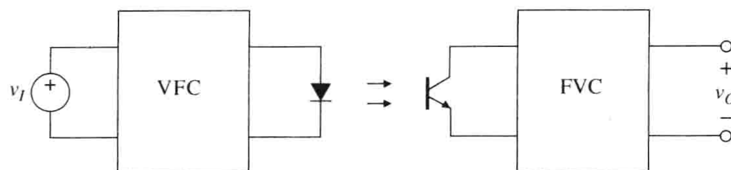


FIGURE 10.37
Transmission of analog information in isolated form.

The value of C is determined on the basis of a maximum duty cycle of 25%, as discussed earlier, while R now establishes the full-scale value of v_O. As in the VFC case, the input offset voltage of *OA* should be nulled to avoid degrading the conversion accuracy at the low end of the range.

Between consecutive closures of *SW*, R will cause C_1 to discharge somewhat, resulting in output ripple. This can be objectionable, especially at the low end of the conversion range where the ripple-to-signal ratio is the worst. The maximum ripple is $V_{r(\max)} = (1 \text{ mA})T_H/C_1$. Using Eq. (10.29), we get

$$V_{r(\max)} = \frac{C}{C_1} 7.5 \text{ V} \tag{10.34}$$

indicating that the ripple can be reduced by making C_1 suitably large. Too large a capacitance, however, slows down the response to a rapid change in f_I since this response is governed by the time constant $\tau = RC_1$. The optimum value of C_1 is, therefore, a compromise between the two opposing demands.

Figure 10.37 shows, in block diagram form, a typical VFC-FVC arrangement for transmitting analog information in isolated form. Here v_I is usually a transducer signal that has been amplified by an instrumentation amplifier. The VFC converts v_I to a train of current pulses for the LED, the phototransistor reconstructs the pulse train at the receiving end, and the FVC converts frequency back to an analog signal v_O. The example shown utilizes an opto-isolator; however, other forms of isolated coupling are possible, such as fiber optic links, pulse transformers, and RF links.

PROBLEMS

10.1 Sine wave generators

10.1 Show that for arbitrary component values in its positive-feedback network, the Wien-bridge circuit of Fig. 10.2*a* gives $B(jf_0) = 1/(1 + R_s/R_p + C_p/C_s)$ and $f_0 = 1/(2\pi\sqrt{R_s R_p C_s C_p})$, where R_p and C_p are the parallel and R_s and C_s the series elements. Hence, verify that neutral stability requires $R_2/R_1 = R_s/R_p + C_p/C_s$.

10.2 Disregarding the limiter in Fig. 10.3*a*, obtain expressions for $T(s)$ for the cases in which the feedback resistance is 22.1 kΩ, 20.0 kΩ, and 18.1 kΩ. Then, find the pole locations for each of the three cases.

10.3 Problem 10.1 indicates that the frequency of a Wien-bridge oscillator can be varied by varying, for instance, R_p. However, to maintain neutral stability, we must also vary R_s in such a way as to keep the ratio R_s/R_p constant. This awkward constraint is avoided

by the circuit[9] of Fig. P10.3. (*a*) Show that f_0 is still as in Problem 10.1, but neutral stability now requires $(R_2/R_1)(1 + R_3/R_p) = R_s/R_p + C_p/C_s$. (*b*) Verify that if we let $R_2/R_1 = C_p/C_s$, this condition simplifies to $R_3 = (R_1/R_2)R_s$. (*c*) Assuming sufficiently fast JFET-input of amps in the design shown, find the range of variability of f_0.

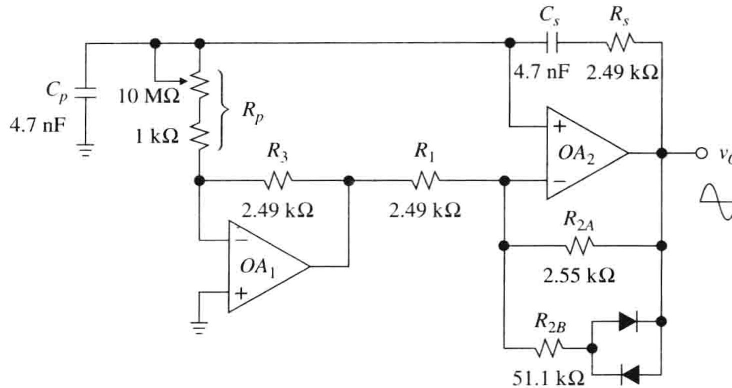


FIGURE P10.3

10.4 Consider a Wien-bridge oscillator whose parallel elements are $C_p = 2.0$ nF and $R_p = 10$ kΩ, and whose series elements are $C_s = 1.0$ nF and $R_s = 20$ kΩ. Assuming ±9-V power supplies, design a nonlinear network of the type of Fig. 10.5 to ensure oscillations of ±5-V peak values. Use PSpice to fine-tune your design. What is the predicted and the actual frequency of oscillation? What happens if the diodes are removed?

10.5 Adapt the multiple-feedback circuit of Fig. P3.41 to design a 100-kHz sinusoidal oscillator with peak values of ±3 V. Assume ±5-V power supplies, and verify with PSpice.

10.6 Suppose the low-pass filter of Fig. 3.23 has $R_1 = 2R_2 = 20$ kΩ and $C_1 = 2C_2 = 2$ nF, and is powered from ±5-V power supplies. Can the circuit be turned into an oscillator? If no, explain why. If yes, show a suitable design, and find f_0.

10.7 Repeat Problem 10.6, but for the *KRC* filter of Fig. 3.27, assuming $R_1 = R_2 = R_3 = 22.6$ kΩ and $C_1 = C_2 = 2$ nF.

10.8 Estimate the s-plane locations of the poles of the quadrature oscillator of Fig. 10.6*a* at power turn-on, before the diodes kick in.

10.2 Multivibrators

10.9 In the circuit of Fig. 10.7*a* let $R = 330$ kΩ, $C = 1$ nF, $R_1 = 10$ kΩ, and $R_2 = 20$ kΩ. Assuming ±15-V supplies, find f_0 and $D(\%)$ if a third resistance $R_3 = 30$ kΩ is connected between the noninverting-input pin of the 301 and the −15-V supply.

10.10 In the circuits of Fig. 10.7*a* let $R_1 = R_2 = 10$ kΩ, and suppose a control source v_I is connected to the noninverting input of the comparator via a 10-kΩ series resistance.

Sketch the modified circuit, and show that it allows for automatic duty-cycle control. What are the expressions for $D(\%)$ and f_0 in terms of v_I? What is the permissible range for v_I?

10.11 In the circuits of Fig. 10.9*a* and Fig. 10.12*a* specify suitable components for $f_0 =$ 100 kHz. The circuits must have provision for the exact adjustment of f_0.

10.12 (*a*) Using a 339 comparator, design a single-supply astable multivibrator with $f_0 =$ 10 kHz and $D(\%) = 60\%$. (*b*) Repeat (*a*), but with $D(\%) = 40\%$.

10.13 The inverters of Fig. 10.12*a* have the following threshold ratings at $V_{DD} = 5$ V: $V_T = 2.5$ V typical, 1.1 V minimum, and 4.0 V maximum. (*a*) Specify suitable components for $f_0 = 100$ kHz typical. (*b*) Find the percentage spread of f_0 due to the spread of V_T.

10.14 Compared to the two-gate oscillator of Fig. 10.12*a*, the three-gate counterpart of Fig. P10.14 is always guaranteed to start. Assuming $V_T = 0.5V_{DD}$, sketch the timing waveforms and derive an expression for f_0.

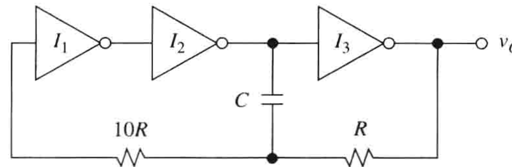


FIGURE P10.14

10.15 If in Fig. P10.14 we remove the capacitor and replace each resistor with a wire, the resulting circuit is called a *ring oscillator* and is often used to measure the propagation delays of logic gates. (*a*) Sketch the voltages at the gate outputs versus time; then derive a relationship between the average gate propagation delay t_P and the frequency of oscillation f_0. (*b*) Can this technique be extended to four gates within the loop? Explain.

10.16 Assuming the threshold spread specifications of Problem 10.13, find suitable components for $T = 10\ \mu$s (typical) in the one-shot of Fig. 10.14*a*; then find the percentage spread of T.

10.17 Design a one-shot using two CMOS NAND gates. Next, explain how it works, show its waveforms, and derive an expression for T. (Recall that the output of a NAND gate goes low only when both inputs are high.)

10.18 Consider the circuit obtained from the one-shot of Fig. 10.14*a* by connecting the output of G to the input of I directly, inserting a resistance R between the lower input of G and ground, and returning the output of I to the lower input of G via a series capacitance C. Draw the modified circuit; then, sketch and label its waveforms, and find T if $R = 100$ kΩ, $C = 220$ pF, and $V_T = 0.4V_{DD}$.

10.3 Monolithic timers

10.19 Let the 555 astable multivibrator of Fig. 10.16*a* be modified as follows: R_B is shorted out, and the wire connecting the bottom node of R_A to pin 7 is cut to allow for the

insertion of a series resistance R_C. (*a*) Sketch the modified circuit and show that choosing $R_C = R_A/2.362$ gives $D(\%) = 50\%$. (*b*) Specify suitable components for $f_0 = 10$ kHz and $D(\%) = 50\%$.

10.20 (*a*) Verify that if the THRESHOLD and TRIGGER terminals of the TLC555 CMOS timer are tied together to form a common input, then the device forms an inverting Schmitt trigger with $V_{TL} = (1/3)V_{DD}$, $V_{TH} = (2/3)V_{DD}$, $V_{OL} = 0$ V, and $V_{OH} = V_{DD}$, where V_{DD} is the supply voltage. (*b*) Using just one resistor and one capacitor, configure the device as a 100-kHz free-running multivibrator, and verify that its duty cycle is 50%.

10.21 Design a 555 one-shot whose pulse width can be varied anywhere from 1 ms to 1 s by means of a 1-MΩ pot.

10.22 A 10-μs 555 one-shot is powered from $V_{CC} = 15$ V. What voltage must be applied to the CONTROL input to stretch T from 10 μs to 20 μs? To shrink T from 10 μs to 5 μs?

10.23 Using a 555 timer powered from $V_{CC} = 5$ V, design a voltage-controlled astable multivibrator whose frequency of oscillation is $f_0 = 10$ kHz when $V_{TH} = (2/3)V_{CC}$, but can be varied over the range 5 kHz $\le f_0 \le$ 20 kHz by externally varying V_{TH}. What are the values of V_{TH} and $D(\%)$ corresponding to the extremes of the above frequency range?

10.24 In the circuit of Fig. 10.18*a* specify suitable components and output interconnections for $T = 1$ s and $T_o = 3$ min.

10.4 Triangular wave generators

10.25 In the circuit of Fig. 10.19*a* let the noninverting input of *OA* be lifted off ground and returned to a +3-V source. Draw the modified circuit; then, sketch and label its waveforms and find f_0 and $D(\%)$ if $R = 30$ kΩ, $C = 1$ nF, $R_1 = 10$ kΩ, $R_2 = 13$ kΩ, $R_3 = 2.2$ kΩ, and D_5 is a 5.1-V reference diode.

10.26 In the circuit of Fig. 10.19*a* let $R_1 = R_2 = R = 10$ kΩ, $R_3 = 3.3$ kΩ, $V_{D(\text{on})} = 0.7$ V, $V_{Z5} = 3.6$ V, and suppose a control source v_I is connected to the inverting input of *OA* via a 10-kΩ series resistance. Sketch the modified circuit, and show that it allows for automatic duty-cycle control. What are the expressions for $D(\%)$ and f_0 in terms of v_I? What is the permissible range for v_I?

10.27 Using an LF411 op amp and an LM311 comparator, both of which have macromodels in PSpice's library, (*a*) design a circuit of the type of Fig. 10.19*a* to produce 100-kHz triangular/square waves with ±5-V peak amplitudes. Assume ±9-V power supplies. (*b*) Verify with PSpice. (*c*) Try different PSpice runs with progressively decreasing values of C until circuit behavior starts to depart appreciably from the intended. How far up in frequency can you push this circuit?

10.28 In the circuit of Fig. 10.20*a* specify suitable components so that both waves have peak amplitudes of 5 V and T_L and T_H are independently adjustable from 50 μs to 50 ms.

10.29 Using a CMOS op amp connected as a Deboo integrator, and a CMOS 555 timer connected as a Schmitt trigger in the manner of Problem 10.20, design a single-supply triangular wave generator. Then, show its waveforms and derive an expression for f_0.

10.30 The effect of component tolerances in the VCO of Fig. 10.21*a* can be compensated for by inserting a variable resistance R_s in series between the control source v_I and the rest of the circuit, and suitably decreasing the nominal value of C to allow for the adjustment of k in both directions. Design a VCO with $k = 1$ kHz/V, k adjustable over a range of $\pm 25\%$.

10.31 Shown in Fig. P10.31 is another popular VCO. Sketch and label its waveforms, and find an expression for f_0 in terms of v_I.

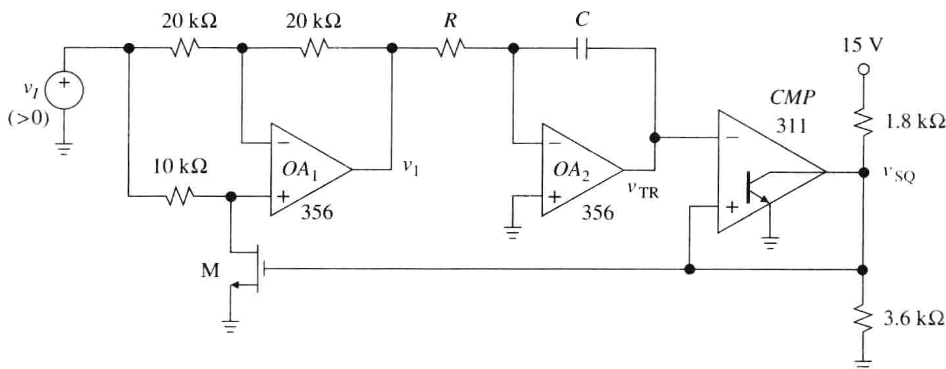


FIGURE P10.31

10.32 Design a wave-shaper circuit that accepts the triangular output of the VCO of Fig. 10.21 and converts it to a sine wave of variable amplitude and offset. Amplitude and offset must be separately adjustable over the ranges 0 to 5 V and -5 V to $+5$ V, respectively.

10.33 Figure P10.33 shows a crude triangular-to-sine wave converter. R_1 and R_2 are found by imposing that v_S and $v_T/(1 + R_2/R_1)$ have (*a*) identical slopes at the zero crossings, and (*b*) peak values equal to $V_{D(\text{on})}$. Assuming $V_{D(\text{on})} = 0.7$ V at $I_D = 1$ mA, find R_1 and R_2 if v_{TR} has peak values of ± 5 V; then use PSpice to plot v_T and v_S versus time.

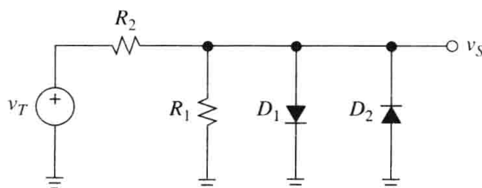


FIGURE P10.33

10.34 The crude triangle-to-sine converter of Fig. P10.33 can be improved considerably by *rounding* the sides of the triangular wave input, besides *clipping* it at the top and bottom. The circuit of Fig. P10.34 provides a VTC with a slope of 1 V/V near the

origin, where all diodes are off. As the magnitude of v_T rises and approaches a diode drop, either D_1 or D_2 goes on, in effect switching R_2 into the circuit. At this point, the slope of the VTC decreases to about $R_1/(R_1 + R_2)$. As the magnitude of v_T rises further and v_S approaches two diode drops, either the D_3-D_4 pair goes on, clipping the top of the waveform, or the D_5-D_6 pair goes on, clipping the bottom. Let us arbitrarily impose $V_{sm} = 2 \times 0.7 = 1.4$ V, so that $V_{tm} = (\pi/2)1.4 = 2.2$ V, and let us assume diodes with $I_s = 2$ fA and $nV_T = 26$ mV, so that at 0.7 V they draw 1 mA. To find suitable values for R_1 and R_2, arbitrarily impose the following pair of constraints: (1) When v_T reaches its positive peak V_{tm}, let the current through the D_3-D_4 pair be 1 mA; (2) When v_T reaches *half* its positive peak, or $V_{tm}/2$, let the slope of the VTC match that of the sine function there, which one can readily prove to be cos $45^\circ = 0.707$ V/V. (*a*) Guided by the above constraints, find suitable values for R_1 and R_2. (*b*) Simulate the circuit via PSpice using a 1-kHz triangular wave with peak values of ± 2.2 V, and display both v_T and v_S versus time. Try different runs, each time changing the values of R_1 and R_2 a bit until you come up with a set that gives what you think is the best sine wave. (*c*) Using the optimized wave shaper of part (*b*) as a basis, design a circuit that accepts a triangular wave with peak values of ± 5 V and yields a sine wave also with peak values of ± 5 V. *Hint:* At the input, replace R_1 by a suitable voltage divider to accommodate the increased triangular wave while still meeting the aforementioned constraints. At the output, use a suitable amplifier implemented with a 741-type op amp.

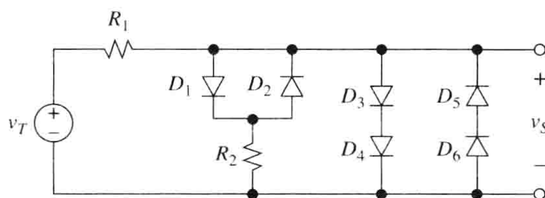


FIGURE P10.34

10.5 Sawtooth wave generators

10.35 (*a*) Show that connecting a resistance R_4 between the source v_I and the inverting-input pin of the 311 in Fig. 10.24*a* gives $V_T = V_{T0} - k|v_I|$, $V_{T0} = V_{CC}/[1 + R_2/(R_3 \parallel R_4)]$, and $k = 1/[1 + R_4/(R_2 \parallel R_3)]$. (*b*) Verify that letting $R_4 = (R_2 \parallel R_3)(RC/T_D - 1)$ gets rid of the T_D term in Eq. (10.21) and gives $f_0 = |v_I|/(RCV_{T0})$. (*c*) Assuming $T_D \cong 0.75\ \mu$s, specify suitable components for a sensitivity of 2 kHz/V and a low-frequency sawtooth amplitude of 5 V. Your circuit is to be compensated against the error due to T_D.

10.6 Monolithic waveform generators

10.36 Derive Eq. (10.23).

10.37 Assuming $V_{CC} = 15$ V, design an ICL8038 sawtooth generator with $f_0 = 1$ kHz and $D(\%) = 99\%$. The circuit must have provision for frequency adjustment over a $\pm 20\%$ range.

10.38 Specify C for a 20-kHz full-scale frequency in the VCO of Fig. 10.29.

10.39 Assuming $V_{CC} = 15$ V, design an XR2206 sawtooth generator with $f_0 = 1$ kHz, $D(\%) = 99\%$, and sawtooth peaks of 5 V and 10 V.

10.7 *V-F* and *F-V* converters

10.40 (*a*) Using the AD537 VFC, design a circuit that accepts a voltage in the range $-10\text{ V} < v_S < 10$ V and converts it to a frequency in the range $0\text{ Hz} < f_O < 20$ kHz. The circuit is to be powered from ±15-V poorly regulated supplies. (*b*) Repeat, but for the case of an input $4\text{ mA} < i_S < 20$ mA and an output range $0 < f_O < 100$ kHz.

10.41 Repeat Example 10.8, but for the Fahrenheit scale.

10.42 The circuit of Fig. P10.42 allows for the VFC32 to work with bipolar inputs. (*a*) Analyze the circuit for both $v_I > 0$ and $v_I < 0$, and find a condition for the resistances that will ensure $f_O = k|v_I|$. (*b*) Specify suitable components for a VFC sensitivity of 10 kHz/V.

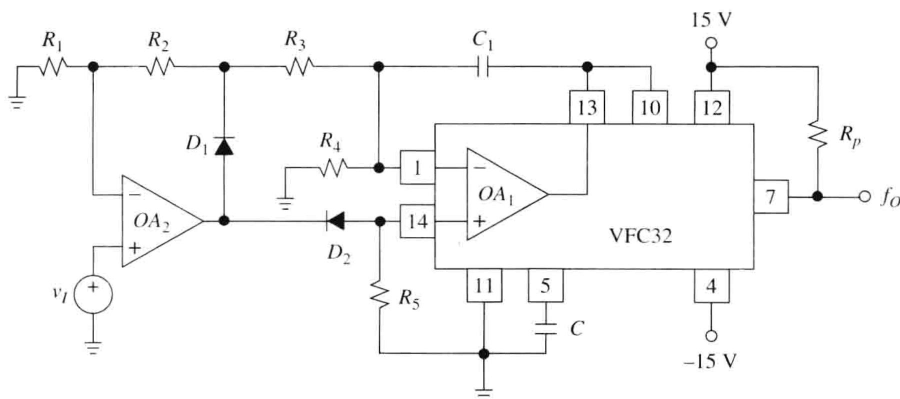


FIGURE P10.42

10.43 Specify suitable component values so that the FVC of Fig. 10.36 yields a full-scale output of 10 V for a full-scale input of 100 kHz with a maximum ripple of 10 mV. Then, estimate how long it takes for the output to settle within 0.1% of the final value for a full-scale change in f_I.

10.44 Using a 4N28 optocoupler, design an external resistive network to provide an opto-coupled link between the VFC of Example 10.9 and the FVC of Problem 10.43. The transistor of the 4N28 gives $I_{C(\text{min})} = 1$ mA with a diode forward current $I_D = 10$ mA. Assume ±15-V supplies.

REFERENCES

1. "Sine Wave Generation Techniques," Texas Instruments Application Note AN-263, http://www.ti.com/lit/an/snoa665c/snoa665c.pdf.
2. E. J. Kennedy, *Operational Amplifier Circuits: Theory and Applications,* Oxford University Press, New York, 1988.

3. J. Williams, "Circuit Techniques for Clock Sources," Linear Technology Application Note AN-12, http://cds.linear.com/docs/en/application-note/an12fa.pdf.
4. A. B. Grebene, *Bipolar and MOS Analog Integrated Circuit Design,* John Wiley & Sons, New York, 1984.
5. *Linear & Telecom ICs for Analog Signal Processing Applications,* Harris Semiconductor, Melbourne, FL, 1993–1994, pp. 7-120–7-129.
6. P. Klonowski, "Analog-to-Digital Conversion Using Voltage-to-Frequency Converters," Analog Devices Application Note AN-276, http://www.analog.com/static/imported-files/application_notes/185321581AN276.pdf.
7. B. Gilbert and D. Grant, "Applications of the AD537 IC Voltage-to-Frequency Converter," Analog Devices Application Note AN-277, http://www.analog.com/static/imported-files/application_notes/511072672AN277.pdf.
8. J. Williams, "Design Techniques Extend V/F Converter Performance," *EDN,* May 16, 1985, pp. 153–164.
9. P. Brokaw, "FET Op Amps Add New Twist to an Old Circuit," *EDN,* June 5, 1974, pp. 75–77.

11

VOLTAGE REFERENCES AND REGULATORS

电压基准 / 稳压电源具有从一个欠稳定的电源 V_I 中得到稳定直流电源 V_O 的功能。它的一般结构如图 11.1 所示。

作为稳压电源的情况，V_I 通常是一个不精确的电源，例如经变压器和二极管简单整流滤波的输出电压。稳压电源的输出 V_O 可以给其他电路供电，这些电路统称为负载，并用从电源中吸取的电流 I_O 来表征。

作为电压基准的情况，V_I 是已经调整到一定程度的电压，电压基准的作用是产生一个更加稳定的电压 V_O，令其为其他电路提供基准。电压基准的作用相当于乐队中确定音高的音叉。例如，数字万用表的满刻度精度就由其内部一个合适质量的电压基准确定。类似地，电源；A-D，D-A，V-F 和 F-V 转换器；传

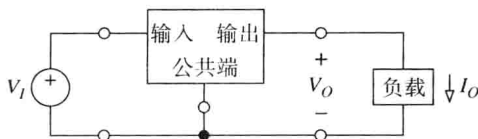


图 11.1 参考电压 / 稳压器的基本连接电路图

感器电路；VCO；对数 / 反对数放大器，以及一些其他电路或系统都需要某种基准或尺度，以满足所需的精度。因此电压基准的主要要求是精确度和稳定性。典型的稳定性要求为 $100 \times 10^{-6}/°C$（每摄氏度百万分之几）或者更好。为了使自身发热引起的误差最小，电压基准应有一定的电流输出能力，通常为几毫安量级。

传统上，电压标准一直由韦斯顿电池确定。韦斯顿电池是一种在 20°C 时产生 1.018636V 可再生电压的化学装置，具有 $40 \times 10^{-6}/°C$ 的温度系数。现在，固态基准已经具备更好的稳定性。尽管半导体器件受温度影响较大，灵活的补偿技术也可以将温度系数保持在 $1 \times 10^{-6}/°C$ 以下！这些技术可得到具有可预期温度系数的电压或电流，能够用于温度检测中。这就构成了各种单片温度传感器和信号调节器的基础。

稳压电源的性能要求与电压基准相同，只是不如对后者那样严格，并且需要大得多的电流输出能力。输出电流范围为 100mA ~ 10A 甚至更高，具体数值由稳压电源的类型决定。

本章讨论两种常见的类型，即线性稳压电源和开关稳压电源。线性稳压电源通过持续调整串联在 V_I 和 V_O 中的功率晶体管来控制 V_O。这种电路结构简单，但因为晶体管的功耗，电路的效率较低。

开关稳压电源中的晶体管是一个高频率的开关，消耗的能量低于持续导通的情况，因此具有更高的效率。另外，与线性稳压电源不同，开关稳压电源可以产生高于输入电压，甚至与输入电压有相反极性的输出电压；它们可以提供相互隔离的多路输出，可以不通过庞大的变压器直接连在交流电源线上。做到这些的代价是需要线圈、电容和更为复杂的控制电路，以及伴随着严重的噪声干扰。即便如此，开关稳压电源还是广泛地应用于功率计算机和便携式仪器中。即便是对于模拟系统的电源设计，也常利用开关稳压电源高效率和轻重量的优点来形成前置稳压电源，再后置线性稳压电源产生净的输出电压。

本章重点

本章最先讨论电压基准和稳压电源的性能指标：线路和负载调整路、纹波

抑制比、温度漂移和电路稳定性，并用熟悉的电路加以评估。

接下来详细讨论电压基准，先从温度补偿齐纳二极管基准开始，再讨论带隙电压基准和单片温度传感器，其中包含一些应用中的变化，比如电流基准设计和热电偶信号调节。

本章接下来讨论线性稳压电源，重点介绍固有保护和实际应用中的温度考虑。同时也描述一些具体供电应用，例如低压差稳压器和供电监督电路。

本章第二部分用于讨论开关稳压电源。先介绍 buck、boost，buck-boost 变换器的基本拓扑，以及线圈、电容的选择和效率的计算。接下来介绍现在广泛应用的电压模式控制（VMC）和峰值电流模式控制（PCMC）。适当详细地介绍 PCMC 下的斜坡补偿，以及升压型变换器中右半平面零点的作用。开关控制的核心是误差放大器设计，这要依据第 8 章的内容。所有这些概念都通过一定数量的设计实例和 PSpice 仿真来实际说明。

11.1 PERFORMANCE SPECIFICATIONS

The ability of a voltage reference or regulator to maintain a constant output under varying external conditions is characterized in terms of performance parameters such as *line* and *load regulation,* and the *thermal coefficient.* In the case of voltage references, *output noise* and *long-term stability* are also significant.

Line and Load Regulation

Line regulation, also called *input,* or *supply regulation,* gives a measure of the circuit's ability to maintain the prescribed output under varying input conditions. In the case of voltage references, the input is typically an unregulated voltage or, at best, a regulated voltage of lower quality than the reference itself. In the case of voltage regulators, the input is usually derived from the 60-Hz line via a step-down transformer, a diode-bridge rectifier, and a capacitor filter and is afflicted by significant ripple. With reference to the symbolism of Fig. 11.1, we define

$$\text{Line regulation} = \frac{\Delta V_O}{\Delta V_I} \qquad (11.1a)$$

where ΔV_O is the output change resulting from a change ΔV_I at the input. Line regulation is expressed in millivolts or microvolts per volt, depending on the case. An alternative definition is

$$\text{Line regulation (\%)} = 100\frac{\Delta V_O/V_O}{\Delta V_I} \qquad (11.1b)$$

with the units being percent per volt. As you consult the catalogs, you will find that both forms are in use.

A related parameter is the *ripple rejection ratio* (RRR), expressed in decibels as

$$\mathrm{RRR_{dB}} = 20\log_{10}\frac{V_{ri}}{V_{ro}} \tag{11.2}$$

where V_{ro} is the output ripple resulting from a ripple V_{ri} at the input. The RRR is used especially in connection with voltage regulators to provide an indication of the amount of ripple (usually 120-Hz ripple) feeding through to the output.

Load regulation gives a measure of the circuit's ability to maintain the prescribed output voltage under varying load conditions, or

$$\text{Load regulation} = \frac{\Delta V_O}{\Delta I_O} \tag{11.3a}$$

Both voltage references and voltage regulators should behave like ideal voltage sources, delivering a prescribed voltage regardless of the load current. The i-v characteristic of such a device is a vertical line positioned at $v = V_O$. A practical reference or regulator exhibits a nonzero output impedance whose effect is a slight dependence of V_O on I_O. This dependence is expressed via the load regulation, in millivolts per milliampere or per ampere, depending on the output current capabilities. The alternative definition

$$\text{Load regulation (\%)} = 100\frac{\Delta V_O/V_O}{\Delta I_O} \tag{11.3b}$$

expresses the above dependence in percent per milliampere or per ampere.

EXAMPLE 11.1. The data sheets of the μA7805 5-V voltage regulator indicate that V_O typically changes by 3 mV when V_I is varied from 7 V to 25 V, and by 5 mV when I_O is varied from 0.25 A to 0.75 A. Moreover, $\mathrm{RRR_{dB}} = 78$ dB at 120 Hz. (*a*) Estimate the typical line and load regulation of this device. What is the output impedance of the regulator? (*b*) Estimate the amount of output ripple V_{ro} for every volt of V_{ri}.

Solution.

(*a*) Line regulation $= \Delta V_O/\Delta V_I = 3 \times 10^{-3}/(25 - 7) = 0.17$ mV/V. Alternatively, line regulation $= 100(0.17$ mV/V$)/(5$ V$) = 0.0033\%$/V. Load regulation $= \Delta V_O/\Delta I_O = 5 \times 10^{-3}/[(750 - 250)10^{-3}] = 10$ mV/A. Alternatively, load regulation $= 100(10$ mV/A$)/(5$ V$) = 0.2\%$/A. The output impedance is $\Delta V_O/\Delta I_O = 0.01\ \Omega$.

(*b*) $V_{ro} = V_{ri}/10^{78/20} = 0.126 \times 10^{-3} \times V_{ri}$. Thus, a 1-V, 120-Hz ripple at the input will result in an output ripple of 0.126 mV.

Thermal Coefficient

The *thermal coefficient* of V_O, denoted as TC(V_O), gives a measure of the circuit's ability to maintain the prescribed output voltage V_O under varying thermal conditions. It is defined in two forms,

$$\mathrm{TC}(V_O) = \frac{\Delta V_O}{\Delta T} \tag{11.4a}$$

in which case it is expressed in millivolts or microvolts per degree Celsius, or

$$\mathrm{TC}(V_O)\ (\%) = 100\frac{\Delta V_O/V_O}{\Delta T} \tag{11.4b}$$

in which case it is expressed in percent per degree Celsius. Replacing 100 by 10^6 gives the TC in parts per million per degree Celsius. Good voltage references have TCs on the order of a few parts per million per degree Celsius.

EXAMPLE 11.2. The data sheets of the REF101KM 10-V precision voltage reference give a typical line regulation of 0.001%/V, a typical load regulation of 0.001%/mA, and a maximum TC of 1 ppm/°C. Find the variation in V_O brought about by: (*a*) a change of V_I from 13.5 V to 35 V; (*b*) a ±10-mA change in I_O; (*c*) a temperature change from 0 °C to 70 °C.

Solution.

(*a*) By Eq. (11.1*b*), 0.001%/V $= 100(\Delta V_O/10)/(35 - 13.5)$, or $\Delta V_O = 2.15$ mV typical.
(*b*) By Eq. (11.3*b*), 0.001%/mA $= 100(\Delta V_O/10)/(\pm 10\text{ mA})$, or $\Delta V_O = \pm 1$ mV typical.
(*c*) By Eq. (11.4*b*), 1 ppm/°C $= 10^6(\Delta V_O/10)/(70\ ^\circ\text{C})$, or $\Delta V_O = 0.7$ mV maximum. You will agree that these are rather small variations for a 10-V source!

In the case of voltage references, *output noise* and *long-term stability* are also important. The data sheets of the aforementioned REF101 give a typical output noise of 6 μV peak-to-peak from 0.1 Hz to 10 Hz, and a typical long-term stability of 50 ppm/(1000 hours). This means that over a period of 1000 hours (about 42 days) the reference output may typically change by $(50 \times 10^{-6})10\text{ V} = 0.5$ mV.

Illustrative Examples

Let us apply the above concepts to the analysis of the classical shunt regulator of Fig. 11.2. The input is a raw voltage assumed to lie within known limits, or $V_{I(\min)} \le V_I \le V_{I(\max)}$. The goal is to produce an output V_O that is as insensitive as possible to both input and load variations. This is achieved by exploiting the nearly vertical *i-v* characteristic of a Zener diode. As depicted in Fig. 11.3*a*, this characteristic can be approximated with a straight line having a slope of $1/r_z$ and

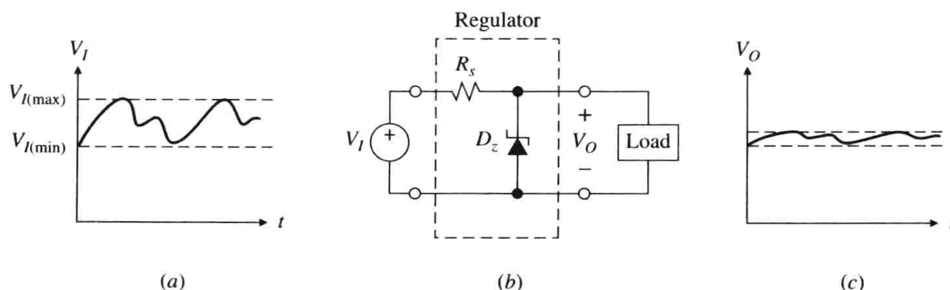


FIGURE 11.2
The Zener diode as a shunt regulator.

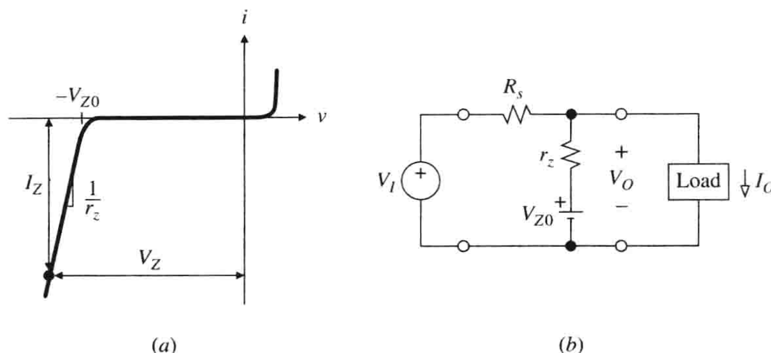


FIGURE 11.3
Breakdown diode characteristic, and equivalent circuit of the shunt regulator.

a v-axis intercept at $-V_{Z0}$, so the coordinates V_Z and I_Z of an arbitrary operating point down the curve are related as $V_Z = V_{Z0} + r_z I_Z$. The resistance r_z, called the *dynamic resistance* of the Zener diode, is typically in the range of a few ohms to several hundreds of ohms, depending on the diode. Zener diodes are specified at the point corresponding to 50% of the power rating. Thus, a 6.8-V, 0.5-W, 10-Ω Zener diode has, at the 50% power point, $I_Z = (P_Z/2)/V_Z = (500/2)/6.8 \cong 37$ mA. Moreover, $V_{Z0} = V_Z - r_z I_Z = 6.8 - 10 \times 37 \times 10^{-3} = 6.43$ V.

It is apparent that a Zener diode can be modeled with a voltage source V_{Z0} and a series resistance r_z, so the circuit of Fig. 11.2*b* can be redrawn as in Fig. 11.3*b*. To function as a regulator, the diode must operate well within the breakdown region under all possible line and load conditions. In particular, I_Z must never be allowed to drop below some safety value $I_{Z(\min)}$. Simple analysis reveals that R_s must satisfy

$$R_s \leq \frac{V_{I(\min)} - V_{Z0} - r_z I_{Z(\min)}}{I_{Z(\min)} + I_{O(\max)}} \tag{11.5}$$

The value of $I_{Z(\min)}$ is chosen as a compromise between the need to ensure proper worst-case operation and the need to avoid excessive power wastage. A reasonable compromise is $I_{Z(\min)} \cong (1/4)I_{O(\max)}$.

We are now ready to estimate the line and load regulation. Applying the superposition principle, we readily find

$$V_O = \frac{r_z}{R_s + r_z} V_I + \frac{R_s}{R_s + r_z} V_{Z0} - (R_s \parallel r_z) I_O \tag{11.6}$$

Only the second term on the right-hand side is a desirable one. The other two indicate dependence on line and load as

$$\text{Line regulation} = \frac{r_z}{R_s + r_z} \tag{11.7a}$$

$$\text{Load regulation} = -(R_s \parallel r_z) \tag{11.7b}$$

Multiplying by $100/V_O$ gives the regulations in percentage form.

EXAMPLE 11.3. A raw voltage $10\text{ V} \le V_I \le 20\text{ V}$ is to be stabilized by a 6.8-V, 0.5-W, 10-Ω Zener diode and is to feed a load with $0 \le I_O \le 10$ mA. (*a*) Find a suitable value for R_s, and estimate the line and load regulation. (*b*) Estimate the effect of full-scale changes of V_I and I_O on V_O.

Solution.

(*a*) Let $I_{Z(\min)} = (1/4)I_{O(\max)} = 2.5$ mA. Then, $R_s \le (10 - 6.43 - 10 \times 0.0025)/(2.5 + 10) = 0.284$ kΩ (use 270 Ω). Line regulation $= 10/(270 + 10) =$ 35.7 mV/V; multiplying by 100/6.5 gives 0.55%/V. Load regulation $= -(10 \parallel 270)$ $= -9.64$ mV/mA, or -0.15%/mA.

(*b*) Changing V_I from 10 V to 20 V gives $\Delta V_O = (35.7\text{ mV/V}) \times (10\text{ V}) = 0.357$ V, which represents a 5.5% change in V_O. Changing I_O from 0 to 10 mA gives $\Delta V_O = -(9.64\text{ mV/mA}) \times (10\text{ mA}) = -0.096$ V, which represents a -1.5% change.

The modest line and load regulation capabilities of a diode can be improved dramatically with the help of an op amp. The circuit of Fig. 11.4 uses the artifice of powering the diode from V_O, that is, from the very voltage we are trying to regulate. The result is a far more stable voltage V_Z, which the op amp then amplifies to give

$$V_O = \left(1 + \frac{R_2}{R_1}\right) V_Z \tag{11.8}$$

This artifice, aptly referred to as *self-regulation,* shifts the burden of line and load regulation from the diode to the op amp. As an additional advantage, V_O is now adjustable, for instance, via R_2. Moreover, R_3 can now be raised to avoid unnecessary power wastage and self-heating effects.

By inspection, we now have

$$\text{Load regulation} \cong -\frac{z_o}{1 + a\beta} \tag{11.9}$$

where a and z_o are the open-loop gain and output impedance, and $\beta = R_1/(R_1+R_2)$. To find the line regulation, we observe that because of single-supply operation, a 1-V change in V_I is perceived by the op amp both as a 1-V supply change and as a 0.5-V input common-mode change. This results in a worst-case input offset voltage change $\Delta V_{OS} = \Delta V_I(1/\text{PSRR} + 1/2\text{CMRR})$ appearing in series with V_Z. The op

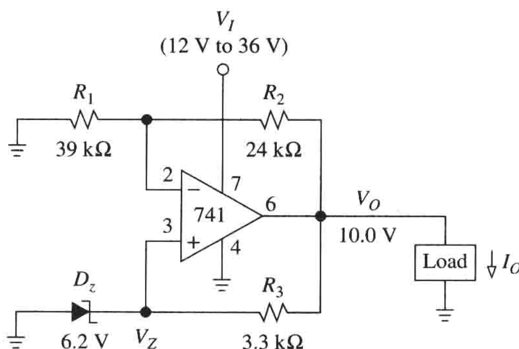


FIGURE 11.4
Self-regulated 10-V reference.

amp then gives $\Delta V_O = (1 + R_2/R_1)\Delta V_{OS}$, so

$$\text{Line regulation} = \left(1 + \frac{R_2}{R_1}\right) \times \left(\frac{1}{\text{PSRR}} + \frac{0.5}{\text{CMRR}}\right) \tag{11.10}$$

We observe that since z_o, a, PSRR, and CMRR are frequency-dependent, so are the line and load regulation. In general, both parameters tend to degrade with frequency.

EXAMPLE 11.4. Assuming typical 741 dc parameters, find the line and load regulation of the circuit of Fig. 11.4.

Solution. Load regulation $= -75/[1 + 2 \times 10^5 \times 39/(39 + 24)] = -0.6\ \mu\text{V/mA} = -0.06$ ppm/mA. Using $1/\text{PSRR} = 30\ \mu\text{V/V}$ and $1/\text{CMRR} = 10^{-90/20} = 31.6\ \mu\text{V/V}$, we get line regulation $= (1 + 24/39) \times (30 + 15.8)10^{-6} = 74\ \mu\text{V/V} = 7.4$ ppm/V. They represent dramatic improvements over the circuit of Example 11.3.

Dropout Voltage

The circuit of Fig. 11.4 will work properly as long as V_I does not drop too low to cause the op amp to saturate. This holds for voltage references and regulators in general, and the minimum difference between V_I and V_O for which the circuit still functions properly is called the *dropout voltage* V_{DO}. In the example of Fig. 11.4 the 741 requires that V_{CC} be at least a couple of volts higher than V_O, so in this case $V_{\text{DO}} \cong 2$ V. Moreover, since the maximum supply rating of the 741 is 36 V, it follows that the permissible input voltage range for the circuit is $12\text{ V} < V_I < 36\text{ V}$.

Start-Up Circuitry

In the self-regulated circuit of Fig. 11.4, V_O depends on V_Z, and V_Z, in turn, depends on V_O being greater than V_Z to keep the diode reverse biased. If at power turn-on V_O fails to swing to a value greater than V_Z, the diode will never turn on, making positive feedback via R_3 prevail over negative feedback via R_2 and R_1. The result is a Schmitt trigger latched in the undesirable state $V_O = V_{OL}$. The possibility for this undesirable behavior is common in most self-biased circuits, and is avoided by using suitable circuitry, known as *start-up circuitry,* to override the amplifier and prevent it from latching in this undesirable state when power is first applied.

The particular implementation of Fig. 11.4 will start properly because of the internal nature of the op amp being used. With reference to Fig. 5.1, we observe that at power turn-on, when v_P and v_N are still zero, the first two stages of the 741 remain off, allowing I_C to turn on the output stage. Consequently, V_O will swing positive until the Zener diode turns on and the circuit stabilizes at $V_O = (1 + R_2/R_1)V_Z$. However, if another op amp type is used, the circuit may never be able to properly bootstrap itself, thus requiring start-up circuitry.

11.2 VOLTAGE REFERENCES

Besides line and load regulation, thermal stability is the most demanding performance requirement of voltage references due to the tendency of IC components to

be strongly influenced by temperature.[2] For example, consider the silicon *pn* junction, which forms the basis of diodes and BJTs. Its forward-bias voltage V_D and current I_D are related as $V_D = V_T \ln(I_D/I_s)$, where V_T is the *thermal voltage* and I_s the *saturation current.* Their expressions are

$$V_T = kT/q \qquad (11.11a)$$

$$I_s = BT^3 \exp(-V_{G0}/V_T) \qquad (11.11b)$$

where $k = 1.381 \times 10^{-23}$ J/K is Boltzmann's constant, $q = 1.602 \times 10^{-19}$ C is the electron charge, T is absolute temperature, B is a proportionality constant, and $V_{G0} =$ 1.205 V is the *bandgap voltage* for silicon.

The TC of the thermal voltage is

$$\text{TC}(V_T) = k/q = 0.0862 \text{ mV/°C} \qquad (11.12)$$

The TC of the junction voltage drop V_D at a given bias I_D is $\text{TC}(V_D) = \partial V_D/\partial T = (\partial V_T/\partial T)\ln(I_D/I_s) + V_T\partial[\ln(I_D/I_s)]/\partial T = V_D/T - V_T\partial(3\ln T - V_{G0}/V_T)/\partial T$. The result is

$$\text{TC}(V_D) = -\left(\frac{V_{G0} - V_D}{T} + \frac{3k}{q}\right) \qquad (11.13)$$

Assuming $V_D = 650$ mV at 25 °C, we get $\text{TC}(V_D) \cong -2.1$ mV/°C. Engineers remember this by saying that the forward drop of a silicon junction decreases by about 2 mV for every degree Celsius increase. Equations (11.12) and (11.13) form the basis of two common approaches to thermal stabilization, namely, *thermally compensated Zener diode references* and *bandgap references.* Equation (11.12) forms also the basis of solid-state *temperature sensors.*

Thermally Compensated Zener Diode References

The thermal stability of V_O in the self-regulated reference of the previous section can be no better than that of V_Z itself. As depicted in Fig. 11.5*a*, $\text{TC}(V_Z)$ is a function of V_Z as well as I_Z. There are two different mechanisms by which the *i-v* characteristic breaks down: *field emission breakdown,* which dominates below about 5 V and produces negative TCs, and *avalanche breakdown,* which dominates above about 5 V and produces positive TCs. The idea behind thermally compensated Zener diodes is to connect a forward-biased diode in series with a Zener diode having an equal but opposing TC, and then fine-tune I_Z to drive the TC of the composite device to zero.[3] This is illustrated in Fig. 11.5*b* for the compensated diodes of the popular 1N821-9 series. The composite device, whose voltage we relabel as $V_Z = 5.5 + 0.7 = 6.2$ V, uses $I_Z = 7.5$ mA to minimize $\text{TC}(V_Z)$. This TC ranges from 100 ppm/°C (1N821) to 5 ppm/°C (1N829).

Self-regulated references based on thermally compensated Zener diodes are available in monolithic form from a number of sources. A popular example is the LT1021 precision reference (look it up on the web), sporting a drift of 5 ppm/°C and a ripple rejection in excess of 100 dB. Another popular device is the LM329 precision reference shown in Fig. 11.6 (bottom). The device uses the Zener diode Q_3 in series

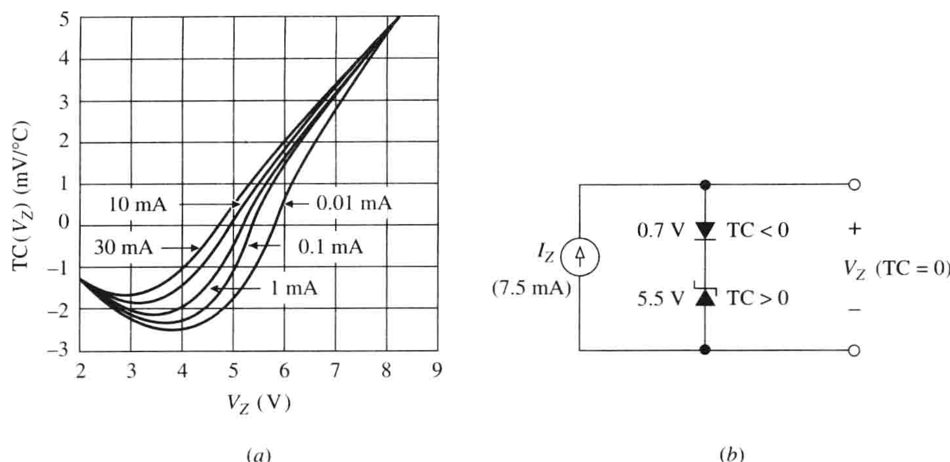


FIGURE 11.5
(*a*) TC(V_Z) as a function of V_Z and I_Z. (*b*) Thermally compensated breakdown diode. (Courtesy of Motorola, Inc.)

with the BE junction of Q_{13} to achieve TCs ranging from 100 ppm/°C to 6 ppm/°C, depending on the version. The device uses also active feedback circuitry to lower the effective dynamic resistance to $r_z = 0.6\ \Omega$ typical, 1 Ω maximum. Except for its much greater stability and much lower dynamic resistance, it acts as an ordinary Zener diode, and it is biased via a series resistor to provide shunt regulation. The bias current may be anywhere between 0.6 mA and 15 mA.

Thermal stability can be improved further via substrate thermostating.[4] The LM399 stabilized reference of Fig. 11.6 uses the aforementioned LM329 active diode (shown at the bottom) to provide the reference proper, and suitable stabilizing circuitry (shown at the top) to sense the substrate temperature and hold it at some set value above the maximum expected ambient temperature. Thermal sensing is done via the BE junction of Q_4, and substrate heating via the power-dissipating transistor Q_1. At power turn-on, Q_1 heats the substrate to 90 °C, where it is then maintained within less than 2 °C over ambient variations from 0 °C to 70 °C. The result is a typical TC of 0.3 ppm/°C. Another thermally stabilized reference is the LTZ1000 ultraprecision reference (look it up online), sporting a thermal drift of 0.05 ppm/°C. An obvious drawback of these devices is the additional power required to heat the chip. For instance, at 25 °C, the LM399 dissipates 300 mW. An LM399 application will be shown in Fig. 11.10.

A notorious problem with breakdown diodes is noise, especially avalanche noise, which plagues devices with breakdown voltages above 5 V, where avalanche breakdown predominates. The use of diode structures of the so-called *buried,* or *subsurface,* type[4] reduces noise significantly while improving long-term stability and reproducibility. The LM399 uses this structure to achieve a typical noise rating of 7 μV (rms) from 10 Hz to 10 kHz. When noise becomes a factor, noise-filtering techniques of the type discussed in Section 7.4 can be used.

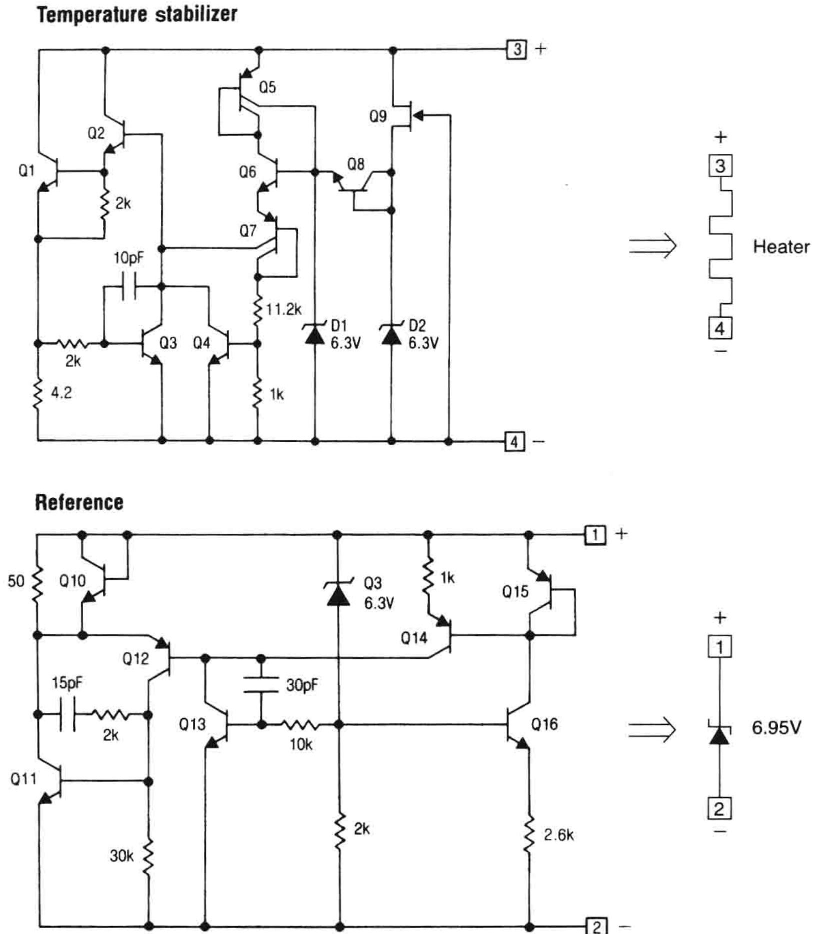


FIGURE 11.6
Circuit diagram of the LM399 6.95-V thermally stabilized reference. (Courtesy of Texas Instruments.)

Bandgap Voltage References

Since the best breakdown voltages range from 6 V to 7 V, they usually require supply voltages on the order of 10 V to operate. This can be a drawback in systems powered from lower supplies, such as 5 V. This limitation is overcome by *bandgap voltage references,* so called because their output is determined primarily by the bandgap voltage $V_{G0} = 1.205$ V. These references are based on the idea of adding the voltage drop V_{BE} of a base-emitter junction, which has a negative TC, to a voltage KV_T proportional to the thermal voltage V_T, which has a positive TC.[2] With reference to Fig. 11.7*a* we have $V_{\text{BG}} = KV_T + V_{BE}$, so $\text{TC}(V_{\text{BG}}) = K\text{TC}(V_T) + \text{TC}(V_{BE})$, indicating that to achieve $\text{TC}(V_{\text{BG}}) = 0$ we need $K = -\text{TC}(V_{BE})/\text{TC}(V_T)$ or, using Eqs. (11.12) and (11.13),

$$K = \frac{V_{G0} - V_{BE}}{V_T} + 3 \tag{11.14}$$

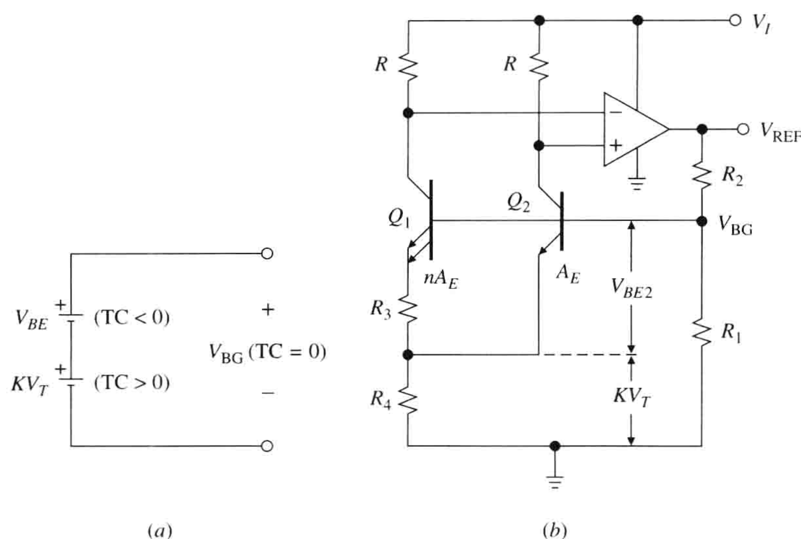


FIGURE 11.7
Bandgap voltage reference.

Substituting into $V_{BG} = KV_T + V_{BE}$ gives

$$V_{BG} = V_{G0} + 3V_T \tag{11.15}$$

At 25 °C we have $V_{BG} = 1.205 + 3 \times 0.0257 = 1.282$ V.

Figure 11.7*b* shows one of several popular bandgap-cell realizations. Known as the *Brokaw cell* for its inventor,[5] the circuit is based on two BJTs of different emitter areas. The emitter area of Q_1 is n times as large as the emitter area A_E of Q_2, so the saturation currents satisfy $I_{s1}/I_{s2} = n$, by Eq. (5.32). With identical collector resistances, the collector currents are also identical, by op amp action. Ignoring base currents, we have $KV_T = R_4(I_{C1} + I_{C2}) = 2R_4I_{C1}$, or

$$KV_T = 2R_4\frac{V_{BE2} - V_{BE1}}{R_3} = \frac{2R_4}{R_3}V_T \ln\frac{I_{C2}I_{s1}}{I_{s2}I_{C1}} = \frac{2R_4}{R_3}V_T \ln n$$

indicating that

$$K = 2\frac{R_4}{R_3}\ln n \tag{11.16}$$

This constant can be fine-tuned by adjusting the ratio R_4/R_3. The op amp raises the cell's voltage to $V_{REF} = (1 + R_2/R_1)V_{BG}$.

EXAMPLE 11.5. Assuming $n = 4$ and $V_{BE2}(25°\text{C}) = 650$ mV in the circuit of Fig. 11.7*b*, specify R_4/R_3 for TC(V_{BG}) = 0 at 25 °C, and R_2/R_1 for $V_{REF} = 5.0$ V.

Solution. By Eq. (11.14), $K = (1.205 - 0.65)/0.0257 + 3 = 24.6$. Then, $R_4/R_3 = K/(2\ln 4) = 8.87$. Moreover, imposing $5.0 = (1 + R_2/R_1)1.282$ gives $R_2/R_1 = 2.9$.

The bandgap concept is available also in CMOS technology, where it is implemented by exploiting the presence of parasitic BJTs such as the so-called well BJT

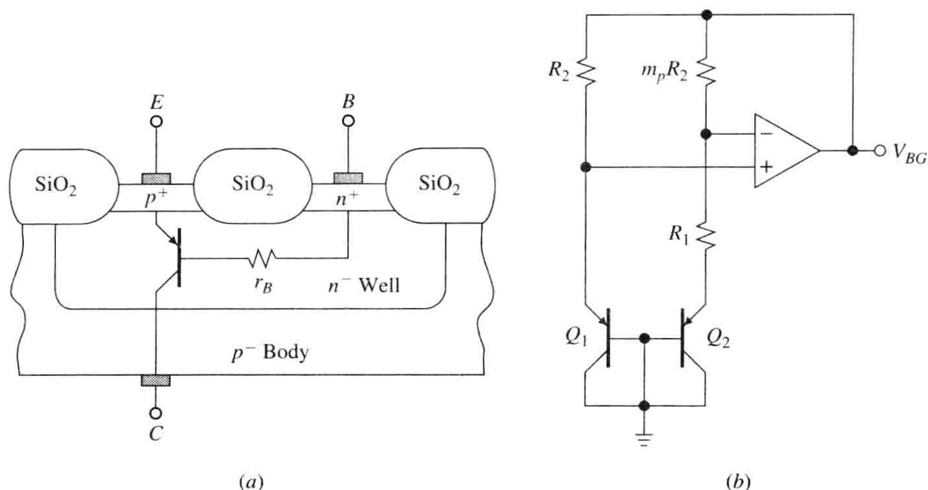


FIGURE 11.8
(*a*) *pnp* BJT fabricated in an n^- CMOS process, and (*b*) CMOS bandgap reference.

depicted in Fig. 11.8*a*. Since the p^- body must be connected to the most negative voltage in the circuit (ground, in this case), the rest of the circuitry must be properly adapted, and Fig. 11.8*b* shows one possibility. The BJTs are usually fabricated with equal emitter areas, so the KV_T term is generated by biasing the BJTs differently via unequal resistances. Through negative feedback, the op amp keeps its own input terminals at the same potential, so Q_2 is forced to operate at a current m_p *times smaller* than Q_1's. It is left as an exercise (see Problem 11.7) to prove that

$$V_{BG} = V_{EB1} + KV_T \qquad K = m_p \frac{R_2}{R_1} \ln(m_p) \tag{11.17}$$

A well BJT tends to exhibit high bulk resistance r_B across the long and lightly doped base region, so in order to minimize the voltage drop across r_B it is customary to bias well BJTs at suitably low currents.

Thanks to their ability to operate with low supply voltages, bandgap references (see also the alternative realizations[2] of Problems 11.5 and 11.6) find wide application as part of systems such as voltage regulators; D-A, A-D, *V-F*, and *F-V* converters; bar graph meters; and power-supply supervisory circuits. They are also available as stand-alone products, either as *two-terminal* or as *three-terminal references,* and sometimes they come with provisions for external trimming.

An example of a two-terminal reference is the already familiar LM385 2.5-V micropower reference diode. Besides the bandgap cell, the device includes circuitry to minimize its dynamic resistance as well as raise the cell voltage to 2.5 V. Typically, it has a TC of 20 ppm/°C and a dynamic resistance of 0.4 Ω. It is biased with a plain series resistance, and its operating current may be anywhere between 20 μA and 20 mA.

An example of a three-terminal reference is the REF05 5-V precision reference. Its output, rated at 5.00 V±30 mV, can be adjusted externally over a ±300-mV range. The REF05A version has, typically, TC = 3 ppm/°C for $-55\ ^\circ\mathrm{C} \le T \le 125\ ^\circ\mathrm{C}$, line

regulation = 0.006%/V for 8 V $\leq V_I \leq$ 33 V, load regulation = 0.005%/mA for $0 \leq I_O \leq 10$ mA, output noise = 10 μV peak-to-peak from 0.1 Hz to 10 Hz, and long-term stability = 65 ppm/1000 hours.

Monolithic Temperature Sensors

The voltage KV_T arising in bandgap cells is linearly *proportional to absolute temperature* (PTAT). As such it forms the basis for a variety of monolithic temperature sensors[6] known as *VPTATs* and *IPTATs,* depending on whether they produce a PTAT voltage or a PTAT current. These sensors enjoy the low-cost advantages of IC fabrication and do not require the costly linearization circuitry common to other sensors, such as thermocouples, RTDs, and thermistors. Besides temperature measurement and control, common applications include fluid-level detection, flow-rate measurement, anemometry, PTAT circuit biasing, and thermocouple cold-junction compensation. Moreover, IPTATs are used in remote-sensing applications because of their insensitivity to voltage drops over long wire runs.

A popular VPTAT is the LM335 precision temperature sensor. As shown in Fig. 11.9*a*, this device acts as a reference diode, except that its voltage is PTAT with TC(V) = 10 mV/K. Thus, at room temperature it gives V(25 °C) = (10 mV/K) × (273.2 + 25)K = 2.982 V. The device is also equipped with a third terminal for the exact adjustment of its TC. The LM335A version comes with an initial room-temperature accuracy of ±1 °C. After calibration at 25 °C, its typical accuracy is ±0.5 °C for −40 °C $\leq T \leq$ 100 °C. Its operating current may be anywhere between 0.5 mA and 5 mA, and its dynamic resistance is less than 1 Ω.

A popular IPTAT is the AD590 two-terminal temperature transducer. To the user this device appears as a high-impedance current source providing 1 μA/K. Terminating it on a grounded resistance as in Fig. 11.9*b* gives a VPTAT with a sensitivity of R × (1 μA/K). The AD590M version comes with a room-temperature accuracy of ±0.5 °C maximum. After calibration at 25 °C, the accuracy is ±0.3 °C maximum for −55 °C $\leq T \leq$ 150 °C. The device operates properly as long as the voltage across its terminals is between 4 V and 30 V.

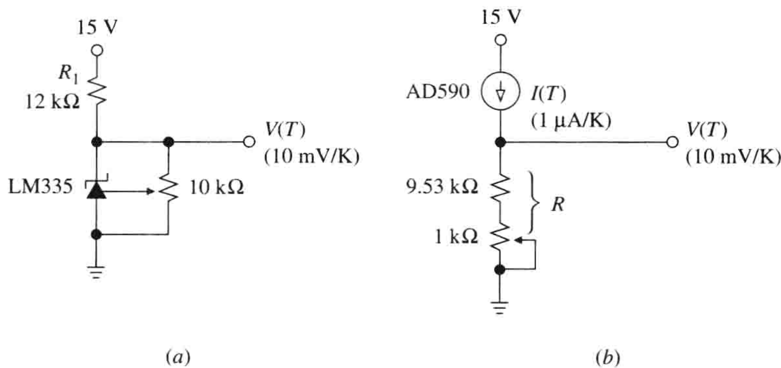


FIGURE 11.9
Basic temperature sensors using the LM335 VPTAT and the AD590 IPTAT.

Additional temperature-processing devices include Celsius and Fahrenheit sensors, and thermocouple signal conditioners. Consult the manufacturer websites to see what is available.

11.3 VOLTAGE-REFERENCE APPLICATIONS

When applying voltage references, care must be exercised to prevent the external circuitry and wiring interconnections from degrading the performance of the reference. This may require the use of precision op amps and low-drift resistors, along with special wiring and circuit-construction techniques. As an example, consider the circuit of Fig. 11.10, which uses a precision op amp to raise the output of a thermally stabilized reference to 10.0 V.

EXAMPLE 11.6. Considering that the LM399 data sheets give $TC_{max} = 2$ ppm/°C and $r_{z(max)} = 1.5\ \Omega$, and the LT1001 data sheets give $TC(V_{OS})_{max} = 1\ \mu V/°C$, $TC(I_B) \cong 4$ pA/°C, $CMRR_{min} = 106$ dB, and $PSRR_{min} = 103$ dB, estimate the worst-case drift and worst-case line regulation of the 10.0-V output.

Solution. The maximum drift due to the LM399 is $2 \times 10^{-6} \times 6.95 = 13.9\ \mu V/°C$, and that due to the overall input error of the LT1001 is $1 \times 10^{-6} + (20 \parallel 8.87)10^3 \times 4 \times 10^{-12} \cong 1\ \mu V/°C$; consequently, the worst-case output drift is $(1 + 8.87/20) \times (13.9 + 1) = 1.44 \times 14.9 = 21.5\ \mu V/°C$. The worst-case line regulation due to the LM399 is $1.5/(1.5 + 7500) = 200\ \mu V/V$, and that due to the LT1001 is $10^{-103/20} + 0.5 \times 10^{-106/20} = (7.1 + 2.5) = 9.6\ \mu V/V$; consequently, the overall worst-case line regulation is $1.44(200 + 9.6) = 303\ \mu V/V$. To give an idea, a 1-V power-supply change has the same effect as a temperature change of $303/21.5 \cong 14$ °C. It is apparent that the use of a precision op amp causes negligible degradation in the present example.

Current Sources

A voltage reference can readily be turned into a current reference[7] by bootstrapping its common terminal with a voltage follower, as in Fig. 11.11. By op amp action, the

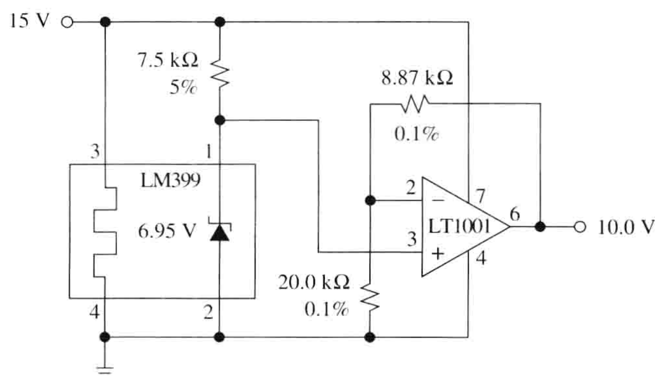


FIGURE 11.10
Buffered 10-V reference.

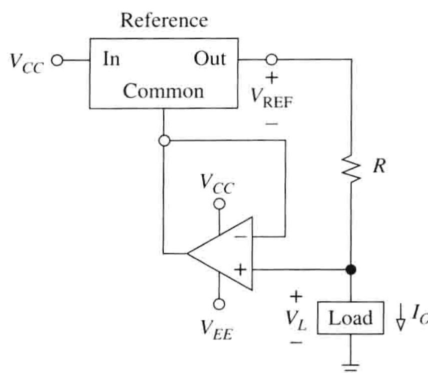


FIGURE 11.11
Turning a voltage reference into a current source.

voltage across R is always V_{REF}, so the circuit gives

$$I_O = \frac{V_{REF}}{R} \tag{11.18}$$

regardless of the voltage V_L developed by the load, provided no saturation effects occur. The permissible range of values of V_L is called the *voltage compliance* of the current source.

EXAMPLE 11.7. Suppose the circuit of Fig. 11.11 uses a 5-V reference with TC = 20 μV/°C, line regulation = 50 μV/V, and dropout voltage V_{DO} = 3 V, and a JFET-input op amp with TC(V_{OS}) = 5 μV/°C and CMRR$_{dB}$ = 100 dB. (*a*) Specify R for I_O = 10 mA. (*b*) Find the worst-case values of TC(I_O) and of the resistance R_o seen by the load. (*c*) Assuming ±15-V supplies, find the voltage compliance.

Solution.

(*a*) $R = 5/10 = 500\ \Omega$ (use 499 Ω, 1%).
(*b*) A 1 °C change in T causes a worst-case change in the voltage across R of $20 + 5 = 25\ \mu$V/°C; the corresponding change in I_O is $25 \times 10^{-6}/500 = 50$ nA/°C. A 1-V change in V_L causes a 50-μV/V change in V_{REF} and a $10^{-100/20} = 10\ \mu$V/V change in V_{OS}, for a worst-case change in I_O of $(50 + 10)10^{-6}/500 = 120$ nA/V. Thus, $R_{o(min)} = (1\text{ V})/(120\text{ nA}) = 8.33\ \text{M}\Omega$.
(*c*) $V_L \le V_{CC} - V_{DO} - V_{REF} = 15 - 3 - 5 = 7$ V.

The bootstrapping principle can readily be applied to the case of diode references to implement either current sources or current sinks. This is shown in Fig. 11.12, where $I_O = V_{REF}/R$ for both circuits. The function of R_1 is to bias the diode. If an LM385 reference diode is used, imposing a bias current of 100 μA when $V_L = 0$ yields $R_1 = 150$ kΩ. The voltage compliance of the source is $V_L \le V_{OH} - V_{REF}$, and that of the sink is $V_L \ge V_{OL} + V_{REF}$. If a 741 op amp and a 2.5-V diode are used, then $V_L \le 10.5$ V for the source, and $V_L \ge -10.5$ V for the sink.

When the circuits just discussed fail to meet load-current demands, we can use current-boosting transistors. The circuit of Fig. 11.13*a* uses a *pnp* BJT to source current. By op amp action, the voltage across the current-setting resistance R is V_{REF}, so the current entering the emitter is $I_E = V_{REF}/R$. The current leaving the

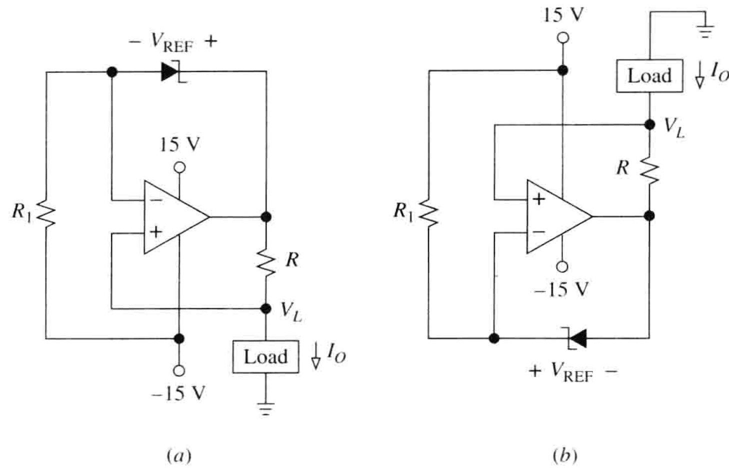


FIGURE 11.12
Using a reference diode to implement a current source and a current sink.

collector is $I_C = [\beta/(\beta + 1)]I_E$, so $I_O = [\beta/(\beta + 1)]V_{REF}/R \cong V_{REF}/R$. The voltage compliance is $V_L \leq V_{CC} - V_{REF} - V_{EC(sat)}$.

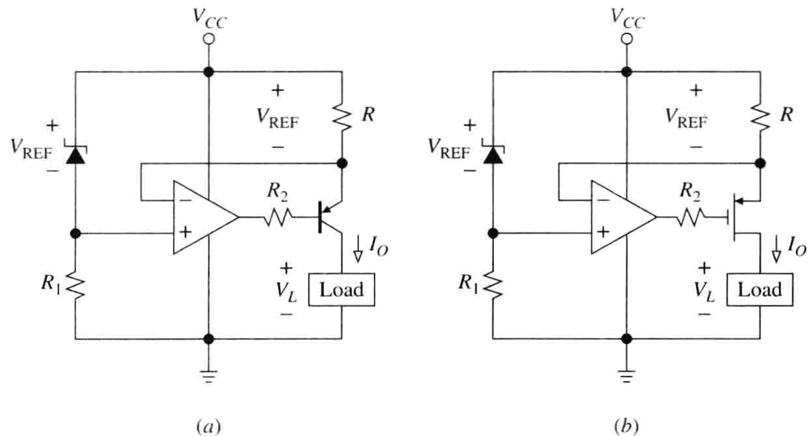


FIGURE 11.13
Current sources with current-boosting transistors.

EXAMPLE 11.8. Let the circuit of Fig. 11.13a use a 741 op amp with $V_{CC} = 15$ V, an LM385 2.5-V diode with a bias current of 0.5 mA, and a 2N2905 BJT with $R_2 = 1$ kΩ. (a) Specify R and R_1 for $I_O = 100$ mA. (b) Assuming typical BJT parameters, find the voltage compliance of the source, and check that the 741 is operating within specifications.

Solution.

(a) We have $R = 2.5/0.1 = 25\ \Omega$ (use 24.9 Ω, 1%), and $R_1 = (15 - 2.5)/0.5 = 25$ kΩ (use 24 kΩ).

(*b*) $V_L \le 15-2.5-0.2 = 12.3$ V. The 741 inputs are at $15-2.5 = 12.5$ V, which is within the input voltage range specifications. Assuming $\beta = 100$ so that $I_B = 1$ mA, we find that the 741 output is at $V_{CC} - V_{\text{REF}} - V_{EB(\text{on})} - R_2 I_B = 15-2.5-0.7-1\times 1 =$ 10.8 V (which is below $V_{OH} = 13$ V), and sinks a current of 1 mA (which is below $I_{sc} = 25$ mA). Consequently, the 741 is operating within specifications.

For higher output currents, the transistor can be replaced by a power *pnp* Darlington, or by a power enhancement *p*-MOSFET as in Fig. 11.13*b*. In these cases, heat-sinking, to be discussed in Section 11.5, may be required.

Temperature-Sensor Applications

In thermometer applications it is desirable that $V(T)$ and $I(T)$ be calibrated in degrees Celsius or Fahrenheit rather than in kelvins. If a VPTAT or an IPTAT is used, then suitable conditioning circuitry is required.[6]

The circuit of Fig. 11.14 senses temperature via the AD590 IPTAT, whose current can be expressed as $I(T) = 273.2\ \mu\text{A} + (1\ \mu\text{A}/^\circ\text{C})T$, T in degrees Celsius. By the superposition principle,

$$V_O(T) = R_2(273.2 + T)10^{-6} - 10R_2/R_1$$

It is apparent that letting $R_1 = 10/(273.2 \times 10^{-6}) = 36.6\ \text{k}\Omega$ will cause a cancellation and leave $V_O(T) = R_2 10^{-6} T$, T in degrees Celsius. For a sensitivity of 100 mV/°C, use $R_2 = (100\ \text{mV})/(1\ \mu\text{A}) = 100\ \text{k}\Omega$. To compensate for the various tolerances, implement R_1 with a 35.7-kΩ resistor in series with a 2-kΩ pot, and R_2 with a 97.6-kΩ resistor in series with a 5-kΩ pot. To calibrate, (*a*) place the IPTAT in an ice bath ($T = 0$ °C) and adjust R_1 for $V_O(T) = 0$ V; (*b*) place the IPTAT in boiling water ($T = 100$ °C) and adjust R_2 for $V_O(T) = 10.0$ V.

Another popular application of temperature sensors is cold-junction compensation in thermocouple measurements.[6] A thermocouple is a temperature sensor consisting of two wires of dissimilar metals and producing a voltage of the type

$$V_{\text{TC}} = \alpha(T_J - T_R)$$

where T_J is the temperature at the measurement or *hot* junction; T_R is the temperature at the reference or *cold* junction, formed where the thermocouple is connected to the leads (usually of copper) of the measuring device; α is the *Seebeck coefficient*. For

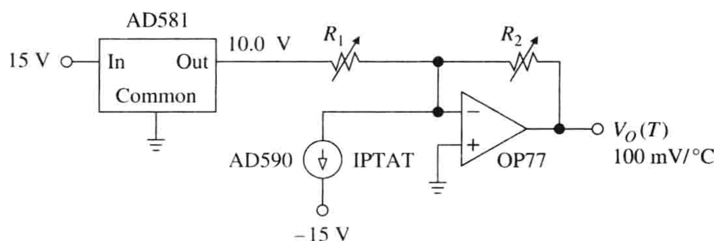


FIGURE 11.14
Celsius sensor.

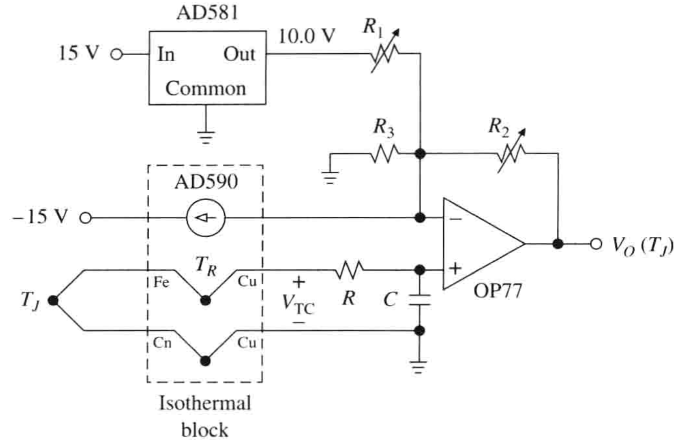


FIGURE 11.15
Thermocouple cold-junction compensation using the AD590 IPTAT.

example, Type J thermocouples are made up of iron and constantan (55% Cu and 45% Ni), and give $\alpha = 52.3\ \mu\text{V}/^\circ\text{C}$.

It is apparent that a thermocouple inherently provides only *relative* temperature information. If we want to measure T_J regardless of T_R, we must use another sensor to measure T_R, as exemplified in Fig. 11.15. Using again the superposition principle,

$$V_O = \left(1 + \frac{R_2}{R_1 \parallel R_3}\right)\alpha(T_J - T_R) + R_2(273.2 + T_R)10^{-6} - 10R_2/R_1$$

where both T_J and T_R are in degrees Celsius. As before, we select R_1 to cancel out the 273.2 term, R_3 to cancel out T_R, and R_2 to achieve the desired output sensitivity.

EXAMPLE 11.9. If the thermocouple of Fig. 11.15 is a type J for which $\alpha = 52.3\ \mu\text{V}/^\circ\text{C}$, specify suitable component values for an output sensitivity of 10 mV/°C. Outline its calibration.

Solution. As before, let $R_1 = 10/(273.2 \times 10^{-6}) = 36.6\ \text{k}\Omega$ to cancel out the 273.2 term. This leaves

$$V_O = \left(1 + \frac{R_2}{R_1 \parallel R_3}\right)\alpha(T_J - T_R) + R_2 T_R 10^{-6}$$

Next, impose $[1 + R_2/(R_1 \parallel R_3)]\alpha = R_2 10^{-6} = 10\ \text{mV}/^\circ\text{C}$ to cancel out T_R as well as achieve the desired output sensitivity. The results are $R_2 = 10.0\ \text{k}\Omega$ and $R_3 = 52.65\ \Omega$.

In practice we would use $R_3 = 52.3\ \Omega$, 1%, and make R_1 and R_2 adjustable as follows: (*a*) place the hot junction in an ice bath and adjust R_1 for $V_O(T_J) = 0$ V; (*b*) place the hot junction in a hot environment of known temperature and adjust R_2 for the desired output (the second adjustment can also be performed with the help of a thermocouple voltage simulator).

To suppress noise pickup by the thermocouple wires, use an *RC* filter as shown, say $R = 10\ \text{k}\Omega$ and $C = 0.1\ \mu\text{F}$.

Thermocouple cold-junction compensators are also available as self-contained IC modules. Two examples are the AD594/5/6/7 series and the LT1025.

11.4 LINEAR REGULATORS

Figure 11.16 shows the basic ingredients of a voltage regulator. The circuit uses the Darlington pair Q_1-Q_2, also called the *series-pass element,* to transfer power from the unregulated input source V_I to the output load at a prescribed regulated voltage V_O. (In MOS technology, the series-pass element is a power FET.) The feedback network R_1-R_2 samples V_O and feeds a portion thereof to the error amplifier EA for comparison against a reference V_{REF}. The amplifier provides the series-pass element with whatever drive it takes to force the error close to zero. The regulator is a classic example of series-shunt feedback, and it can be viewed as a noninverting op amp that has been equipped with a Darlington current booster to give

$$V_O = \left(1 + \frac{R_2}{R_1}\right) V_{REF} \tag{11.19}$$

Since the error amplifier provides currents on the order of milliamperes and the load may draw currents on the order of amperes, a current gain on the order of 10^3 A/A is required. A single power BJT is usually insufficient, so a Darlington pair is used instead, whose overall current gain is $\beta \cong \beta_1 \times \beta_2$. We observe that for an *npn* BJT to work in the forward-active region, where $I_C = \beta I_B$, the conditions $v_{BE} = V_{BE(\text{on})}$ and $v_{CE} \geq V_{CE(\text{sat})}$ must hold. A low-power BJT has typically $\beta \cong 100$, $V_{BE(\text{on})} \cong 0.7$ V, and $V_{CE(\text{sat})} \cong 0.1$ V; a power BJT may have $\beta \cong 20$, $V_{BE(\text{on})} \cong 1$ V, and $V_{CE(\text{sat})} \cong 0.25$ V. If a power MOSFET is used as the series-pass element, then just one transistor suffices because the gate terminal draws virtually zero current.

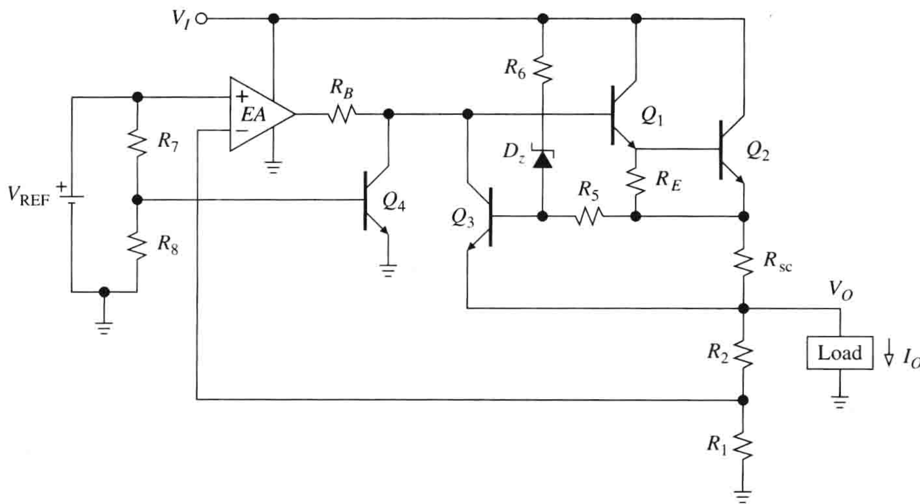


FIGURE 11.16
Simplified circuit diagram of a typical bipolar positive voltage regulator.

Protections

The reliable performance of a power BJT is critically affected by power-dissipation capabilities, current and voltage ratings, maximum junction temperature, and second breakdown, a phenomenon resulting from the formation of hot spots within the BJT, which cause uneven sharing of the total load among different regions of the device. The above factors define a restricted region of the i_C-v_{CE} characteristic, known as the *safe operating area* (SOA), within which the device can be operated without the risk of failure or performance degradation.

Voltage regulators are equipped with special circuitry to protect the power stage against *current overload, second breakdown,* and *thermal overload.* Each circuit is designed to be inactive under normal operating conditions, but to become active as soon as an attempt is made to exceed the corresponding safety limits.

Current overload protection is dictated by maximum power-rating considerations. Since the power dissipated by the series-pass BJT is $P \cong (V_I - V_O)I_O$, we must ensure $I_O \leq P_{\max}/(V_I - V_O)$ for safe operation. The protection scheme of Fig. 11.17, similar to that discussed at the end of Chapter 5 for op amps, uses a brute-force approach to keep I_O below the limit $I_{sc} = P_{\max}/V_I$, which occurs when the output is short-circuited to ground, or $V_O = 0$. As we know, the design equation is

$$R_{sc} = \frac{V_{BE3(\text{on})}}{I_{sc}} \tag{11.20}$$

To confine the series-pass BJT within its SOA, its collector current must be reduced in case the collector-emitter voltage rises above a safety level, a likely event when high-voltage transients are present on the unregulated input line. This protection is implemented with the Zener diode D_z, as shown in Fig. 11.16. This diode, normally in cutoff, is designed to turn on as soon as V_I rises above a safety level. The current supplied by D_z will then turn on Q_3 and divert current away from the base of the series-pass BJT, as in the case of current overload. The function of R_5 is to decouple the base of Q_3 from the low-impedance emitter of Q_2, and that of R_6 is to limit the current through D_z, particularly in the presence of large noise spikes on the input line.

Excessive self-heating may cause permanent damage to BJTs, unless junction temperatures are kept from rising above a safety level, usually 175 °C or less. The

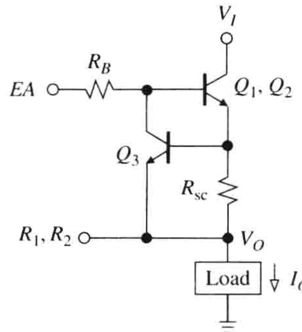


FIGURE 11.17
Output overload protection.

series-pass BJT is protected by sensing its instantaneous temperature and reducing its collector current in case of thermal overload. In the circuit of Fig. 11.16 this protection is provided by Q_4, a BJT mounted in close thermal coupling with the series-pass element. Temperature is sensed by exploiting the negative TC of V_{BE4}. This BJT is designed to be in cutoff during acceptable thermal conditions, but to turn on as soon as the temperature approaches 175 °C. Once in conduction, Q_4 will divert current away from the base of the series-pass BJT, reducing its conduction to the point of even shutting it off until the temperature drops to a more tolerable level.

EXAMPLE 11.10. In the circuit of Fig. 11.16 assume $V_{BE4}(25\ °\text{C}) = 700$ mV and $\text{TC}(V_{BE4}) = -2$ mV/°C. Find R_7 and R_8 to cause thermal shutdown at 175 °C if V_{REF} is a bandgap reference.

Solution. The voltage required to turn on Q_4 can be estimated as $V_{BE4}\ (175\ °\text{C}) = V_{BE4}\ (25\ °\text{C}) + \text{TC}(V_{BE4})(175 - 25)\ °\text{C} \cong 700\ \text{mV} + (-2\ \text{mV}/°\text{C})150\ °\text{C} \cong 400\ \text{mV}$. Ignoring I_{B4} and imposing $0.4 = [R_8/(R_8 + R_7)]1.282$ gives $R_7/R_8 = 2.2$. Assuming $I_{B4} = 0.1$ mA and imposing $V_{\text{REF}}/(R_7 + R_8) \cong 10 I_{B4}$ gives $R_7 = 880\ \Omega$ and $R_8 = 400\ \Omega$.

Efficiency

The efficiency of a regulator is defined as $\eta(\%) = 100 P_O/P_I$, where $P_O (= V_O I_O)$ is the average power delivered to the load, and $P_I (= V_I I_I)$ is that absorbed from the input source. The current I_I drawn from the V_I source splits between the load and the control circuitry consisting of the bandgap reference, the error amplifier, and the feedback network. Being only on the order of milliamperes, the latter current is negligible compared to I_O, so we approximate $I_I \cong I_O$ and write

$$\eta(\%) \cong 100 \frac{V_O}{V_I} \tag{11.21}$$

EXAMPLE 11.11. In the circuit of Fig. 11.16 let $R_B = 510\ \Omega$, $R_E = 3.3\ \text{k}\Omega$, and $R_{\text{sc}} = 0.3\ \Omega$, and suppose the error amplifier can swing its output linearly to within 0.25 V of V_I. Assuming a bandgap reference and typical BJT, find (*a*) the ratio R_2/R_1 for $V_O = 5.0$ V, (*b*) the error-amplifier output voltage and current needed to provide $I_O = 1$ A, (*c*) the dropout voltage V_{DO}, and (*d*) the maximum efficiency for the given I_O. (*e*) What is the efficiency if V_I is obtained from a car battery?

Solution.

(*a*) Imposing the condition $5 = (1 + R_2/R_1) \times 1.282$ gives $R_2/R_1 = 2.9$.
(*b*) For $I_O = 1$ A we have $I_{B2} = I_{E2}/(\beta_2 + 1) = 1/21 = 47.6$ mA, and $I_{E1} = I_{B2} + V_{BE2(\text{on})}/R_{E2} \cong 48$ mA. The error amplifier must therefore supply $I_{OA} = I_{B1} = I_{E1}/(\beta_1 + 1) = 48/101 \cong 0.475$ mA at a voltage of $V_{OA} = V_{R_B} + V_{BE1(\text{on})} + V_{BE2(\text{on})} + V_{R_{\text{sc}}} + V_O = 0.51 \times 0.475 + 0.7 + 1 + 0.3 \times 1 + 5 \cong 7.25$ V.
(*c*) Allowing a headroom of 0.25 V for the EA, it is apparent that to function properly the circuit needs $V_I \geq 7.25 + 0.25 = 7.5$ V. Hence, $V_{\text{DO}} = 7.5 - 5 = 2.5$ V.
(*d*) Since $V_I \geq 7.5$, Eq. (11.21) gives $\eta(\%) \leq 100 \times 5/7.5 \cong 67\%$.
(*e*) For $V_I = 12$ V, the efficiency drops to $\eta(\%) = 100 \times 5/12 \cong 42\%$.

Monolithic Voltage Regulators

The basic structure of Fig. 11.16, or variants thereof, is available in monolithic form from many sources. Two of the earliest products to gain widespread popularity were the μA7800 series of *positive* regulators, and the μA7900 series of *negative* regulators (look them up on the web). The μA78G series is similar to the μA7800 series, except that the R_1-R_2 resistor pair of Fig. 11.16 is omitted, and the inverting input terminal of the error amplifier, referred to as the *control pin,* is made accessible to the user for the external setting of V_O. Called a *four-terminal adjustable regulator,* the device is especially useful in remote sensing. As depicted in Fig. 11.18, mounting the feedback network right across the load and equipping it with separate returns will ensure a regulated voltage of Eq. (11.19) *right at the load,* irrespective of any voltage drops across the stray resistances r_s of the wires. The four-terminal version of the 7900 negative regulators is called the μA79G.

Another popular class of products is offered by *three-terminal adjustable regulators,* of which the LM317 positive regulator and the LM337 negative regulator are among the most widely known examples. In the LM317 functional diagram[8] of Fig. 11.19*a*, the diode is a 1.25-V bandgap reference biased at 50 μA. The error amplifier provides whatever drive it takes to keep the voltage at the output pin 1.25-V higher than the voltage at the adjustment pin. Thus, connecting the device as in Fig. 11.19*b* gives $V_O = V_{\text{ADJ}} + 1.25$ V. By the superposition principle, $V_{\text{ADJ}} = V_O/(1 + R_1/R_2) + (R_1 \parallel R_2)(50\ \mu\text{A})$. Eliminating V_{ADJ} gives

$$V_O = \left(1 + \frac{R_2}{R_1}\right) 1.25\ \text{V} + R_2(50\ \mu\text{A}) \tag{11.22}$$

The purpose of R_1 and R_2, besides setting the value of V_O, is to provide a conductive path toward ground for the quiescent current of the error amplifier and the remaining circuitry in the absence of a load. The data sheets recommend imposing a current of 5 mA through R_1 to meet this requirement. One can then verify that the effect of the 50-μA current becomes negligible, so $V_O = (1 + R_2/R_1)1.25$ V. By varying R_2, V_O can be adjusted anywhere from 1.25 V and 35 V. A more recent adjustable regulator is the LT3080, which you can search online for its circuit schematic as well as useful application suggestions.

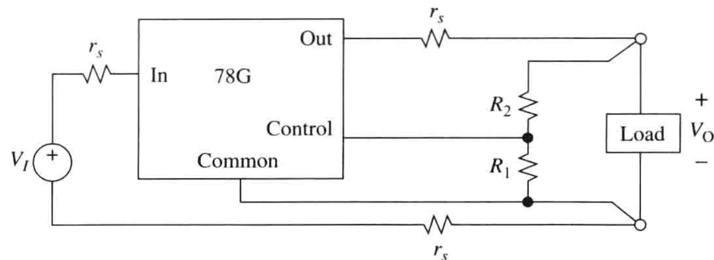


FIGURE 11.18
Adjustable regulator with remote sensing.

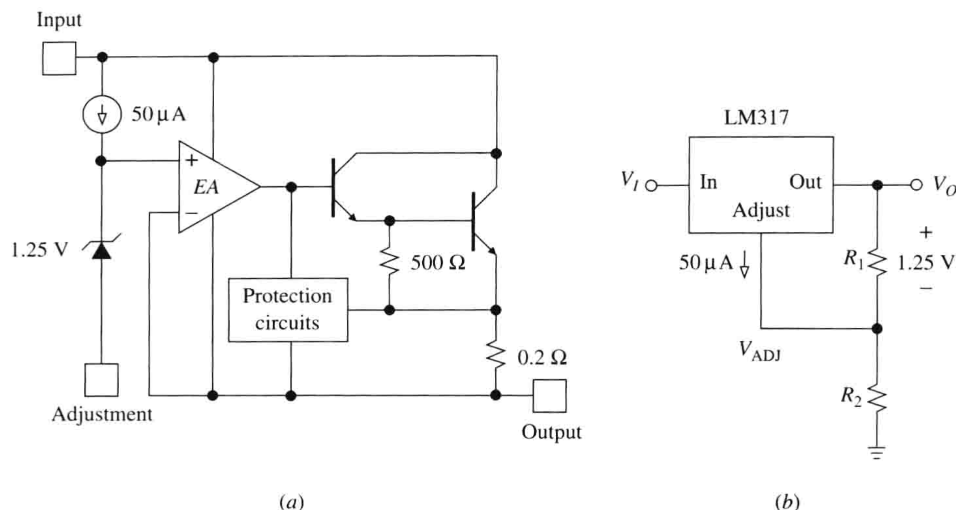


FIGURE 11.19
Functional diagram and typical connection of the LM317 three-terminal adjustable regulator. (Courtesy of Texas Instruments.)

Low-Dropout (LDO) Regulators

Considerations of efficiency and power dissipation, along with the demands by low-voltage portable-power systems, have led to the emergence of low-dropout (LDO) regulators. For a closer look at the dropout voltage $V_{\rm DO}$, refer to the commonly used topologies of Fig. 11.20, each circuit consisting of a series-pass element and its driver, in turn controlled by the voltage v_X. Recalling that $V_{\rm DO} = V_{I(\rm min)} - V_O$, where $V_{I(\rm min)}$ is the lowest input voltage before the control loop stops regulating, we can proceed along the same lines as for the OVS estimations of Section 5.7.

The topology of Fig. 11.20*a* gives $V_{\rm DO} = V_{EC1(\rm EOS)} + V_{BE2(\rm on)} + V_{BE3(\rm on)}$, where $V_{EC1(\rm EOS)}$ is Q_1's emitter-collector voltage at the edge of saturation. Using typical values, $V_{\rm DO} = 0.25 + 0.7 + 1 \cong 2$ V. Figure 11.20*b* shows the MOS version

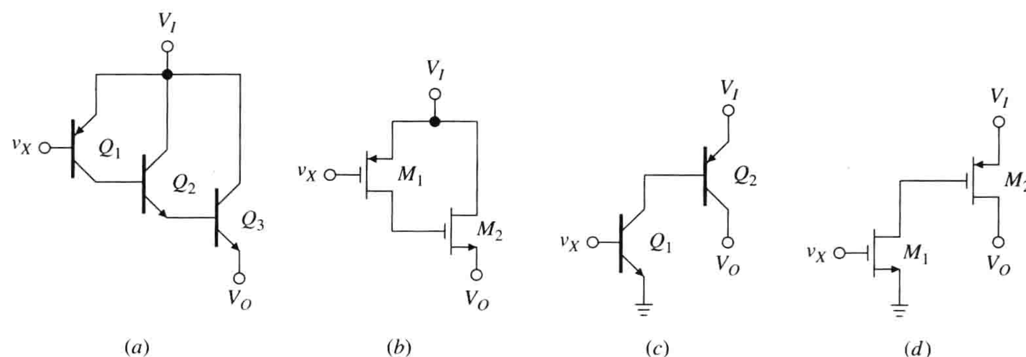


FIGURE 11.20
Common output topologies for series regulators.

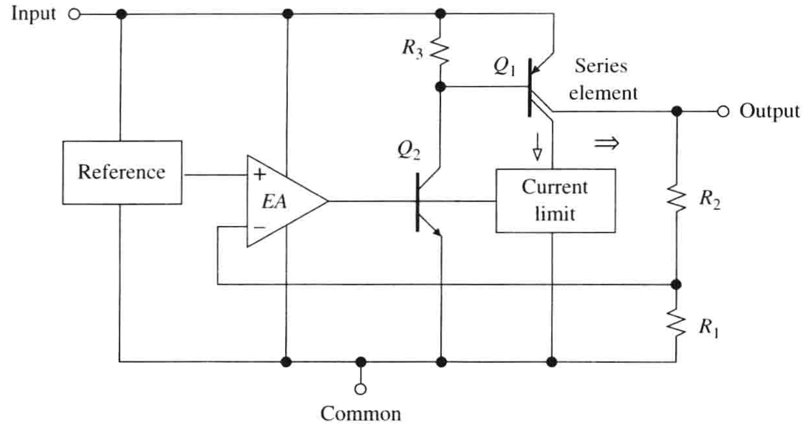


FIGURE 11.21
Block-diagram of a low-dropout (LDO) regulator.

of Fig. 11.20*a* (thanks to the virtually infinite input resistance of FETs, the series-pass function in this case performed by a single power FET, M_2). We now have $V_{DO} = V_{SD1(EOS)} + V_{GS2} = V_{OV1} + V_{tn} + V_{OV2}$, where V_{OV} is the overdrive voltage and V_{tn} is the threshold voltage. Assuming the typical values of Section 5.7, $V_{DO} = 0.25 + 0.75 + 0.25 = 1.25$ V. In both topologies the series-pass element is a voltage follower (common collector in the BJT case, common drain in the MOS case). As such, it is inherently fast, and its low output impedance makes it relatively immune to capacitive loading.[1] However, their V_{DO}s are unacceptably high in many situations.

Following the rail-to-rail considerations of Section 5.7, we can reduce V_{DO} significantly by operating the series-pass element as common emitter (CE) or a common source (CS) instead of CC or CD. This leads to the topologies of Fig. 11.20*c* and *d*, for which we have, respectively, $V_{DO} = V_{EC2(EOS)}$ ($\cong 0.25$ V) and $V_{DO} = V_{SD2(EOS)} = V_{OV2}$ ($\cong 0.25$ V), which are appreciably lower. Their lower V_{DO}s, however, come at a price, and this is the inherently high output impedance of the CE/CS configurations: the feedback pole formed by the capacitive and resistive components of the load tend to destabilize the control loop. A common cure is to use a snubber compensation technique of the type discussed in Section 8.4, using an output capacitor with a prescribed range of capacitance and equivalent series resistance (ESR) values.[1,9]

Figure 11.21 shows an example of a LDO regulator. To avoid using R_{sc}, which would increase V_{DO}, the *pnp* BJT is equipped with an additional small-area collector to provide collector-current sensing information for the overload protection circuitry. LDOs are often used to provide postregulation of the noisier outputs of switching regulators.

11.5
LINEAR-REGULATOR APPLICATIONS

The primary application of voltage regulators is in power supplies, especially distributed supplies, where the unregulated voltage is brought to different subsystems to be treated locally by dedicated regulators. Aside from a few simple requirements,

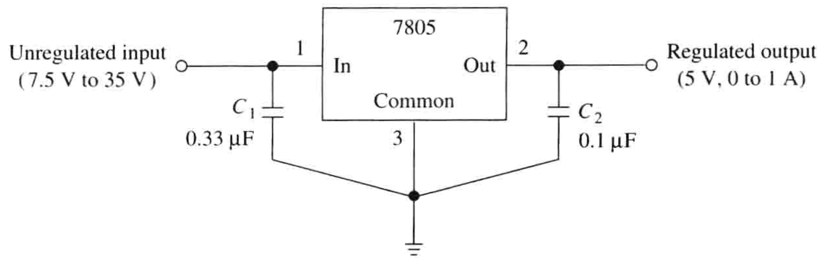


FIGURE 11.22
Typical circuit connection of the μA7805 voltage regulator. (Copyright, Fairchild Semiconductor Corporation, 1982. Used by permission.)

a linear regulator is generally easy to use. As exemplified in Fig. 11.22, the device should always be equipped with an input capacitor to reduce the effects of stray inductance in the input wires, especially if the regulator is located away from the unregulated source, and an output capacitor to help improve the response to sudden load-current changes. For best results, use thick wires and traces, keep the leads short, and mount both capacitors as close as possible to the regulator. Depending on the case, heat-sinking may be required to keep the internal temperature within tolerable levels.

Power Sources

With the help of a few external components, a voltage regulator can, like a voltage reference, be configured for a variety of voltage source or current source applications, the main difference lying in the much higher currents available.

A regulator is configured for a higher output voltage by raising its common terminal to a suitable voltage pedestal. In Fig. 11.23*a* we have $V_O = V_{\text{REG}} + R_2 \times V_O/(R_1 + R_2)$, or

$$V_O = \left(1 + \frac{R_2}{R_1}\right) V_{\text{REG}} \tag{11.23}$$

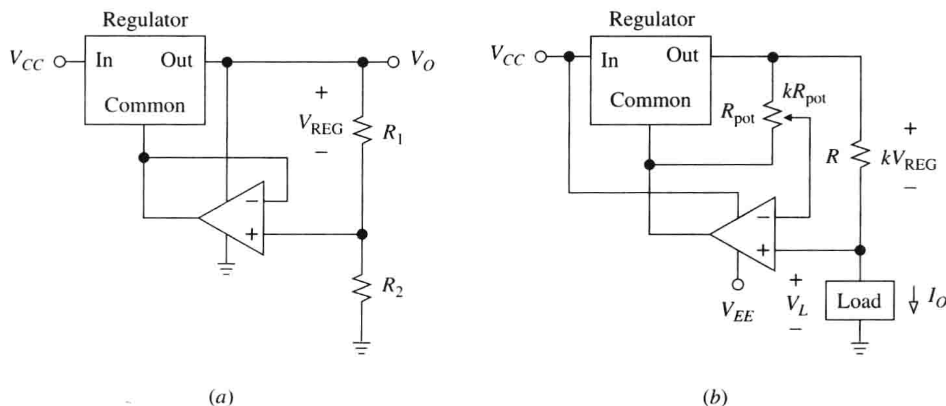


FIGURE 11.23
Configuring a regulator (*a*) as a power voltage source, and (*b*) as an adjustable power current source.

The role of the op amp, which is powered from the regulated output to eliminate any PSRR and CMRR errors, is to prevent the feedback network from being loaded by the common terminal. However, if the current of this terminal is sufficiently small, as in the case of adjustable regulators such as the LM317 and LM337 types, then we can do without the op amp and the circuit simplifies to the familiar form of Fig. 11.19*b*.

EXAMPLE 11.12. Suppose the circuit of Fig. 11.23*a* uses a 7805 5-V regulator, whose ratings are $V_{DO} = 2$ V and $(V_{In} - V_{Common})_{max} = 35$ V. Specify suitable resistances for $V_O = 15.0$ V, and comment on the line and load regulation. What is the permissible range for V_{CC}?

Solution. Imposing $15 = (1 + R_2/R_1)5$ gives $R_2/R_1 = 2$. Use $R_1 = 10$ kΩ, $R_2 = 20$ kΩ. For the exact adjustment of V_O, interpose a 1-kΩ potentiometer between R_1 and R_2, and connect the op amp's noninverting input to the wiper. The percentage values of the line and load regulation are the same as those reported in the 7805 data sheets. However, their mV/V and mV/A values are now $1 + R_2/R_1 = 3$ times as large. The op amp keeps V_{Common} at 10 V, so $V_{CC(max)} = 10 + 35 = 45$ V. Moreover, $V_{CC(min)} = V_O + V_{DO} = 15 + 2 = 17$ V.

In Fig. 11.23*b* the op amp bootstraps the regulator's common terminal with the voltage V_L developed by the output load, and the regulator keeps the voltage across R at kV_{REG}, where k represents the fraction of the potentiometer between the wiper and the regulator's output, $0 \le k \le 1$. Consequently, the circuit gives

$$I_O = k\frac{V_{REG}}{R} \tag{11.24}$$

regardless of V_L, provided no saturation effects occur. We thus have an adjustable current source, and its voltage compliance is $V_L \le V_{CC} - V_{DO} - kV_{REG}$. If a current sink is needed, then we can use a negative regulator. To maximize the compliance for a given V_{CC}, use a regulator with low V_{DO} and V_{REG}. An adjustable regulator of the 317 or 337 type is a good choice.

EXAMPLE 11.13. The circuit of Fig. 11.23*b* uses an LM317 1.25-V regulator, whose ratings are $V_{DO} = 2$ V and line regulation = 0.07%/V maximum. Assuming a 10-kΩ potentiometer, an op amp with $\text{CMRR}_{\text{dB}} \ge 70$ dB, and ±15-V supplies, specify R for an adjustable current from 0 to 1 A; next, find the voltage compliance and the minimum equivalent resistance seen by the load for the case $k = 1$.

Solution. $R = 1.25$ Ω, 1.25 W (use 1.24 Ω, 2 W). $V_L \le 15 - 2 - 1.25 = 11.75$ V. A 1-V change in V_L causes a worst-case change in I_O of $(1.25 \times 0.07/100 + 10^{-70/20})/1.25 = 0.953$ mA, so $R_{o(min)} = (1\text{ V})/(0.953\text{ mA}) = 1.05$ kΩ.

Thermal Considerations

The power dissipated in the base-collector junction of the series-pass BJT is converted into heat, which raises the junction temperature T_J. To prevent permanent damage to the BJT, T_J must be kept within a safe limit. For silicon devices, this limit is in the range of 150 °C to 200 °C. To avoid excessive temperature buildup, heat must be expelled from the silicon chip to the surrounding package structure

and from there to the ambient. At thermal equilibrium, the temperature rise of a constant-power dissipating BJT with respect to the ambient can be expressed as

$$T_J - T_A = \theta_{JA} P_D \tag{11.25}$$

where T_J and T_A are the junction and ambient temperatures, P_D is the dissipated power, and θ_{JA} is the *junction-to-ambient thermal resistance,* in degrees Celsius per watt. This resistance, representing the amount of temperature rise per unit of dissipated power, is given in the data sheets. For instance, for $\theta_{JA} = 50\ °\mathrm{C/W}$ the chip temperature rises above the ambient temperature by 50 °C for every watt of dissipated power. If $T_A = 25\ °\mathrm{C}$ and $P_D = 2$ W, then $T_J = T_A + \theta_{JA} P_D = 25 + 50 \times 2 = 125\ °\mathrm{C}$. We can also regard θ_{JA} as a measure of a device's ability to expel heat. The lower θ_{JA}, the smaller the temperature rise for a given P_D. It is apparent that θ_{JA} and $T_{J(\max)}$ set an upper limit on P_D for a given $T_{A(\max)}$.

The heat-transfer process can be modeled with an electrical-conduction analog where power corresponds to current, temperature to voltage, and thermal resistance to ohmic resistance. This analogy is illustrated in Fig. 11.24 for the case of free-air operation, that is, with no provisions for cooling. The thermal resistance θ_{JA} consists of two components,

$$\theta_{JA} = \theta_{JC} + \theta_{CA} \tag{11.26}$$

where θ_{JC} is the thermal resistance from *junction to case,* and θ_{CA} from *case to ambient.* Using Ohm's law and KVL, we can find the temperature at any point of the heat-flow path once the other parameters are known. If the path involves more than one resistance, the net resistance is the sum of the individual resistances.

The component θ_{JC} is set by device layout and packaging. To help reduce θ_{JC}, the device is encapsulated in a suitably large case, and the collector region, where most of the heat is dissipated, is placed in direct contact with the case. Figure 11.25 shows two popular packages, along with their thermal ratings for the case of the μA7800 and μA7900 series. Data sheets usually give only θ_{JC} and θ_{JA}; then, we can compute $\theta_{CA} = \theta_{JA} - \theta_{JC}$.

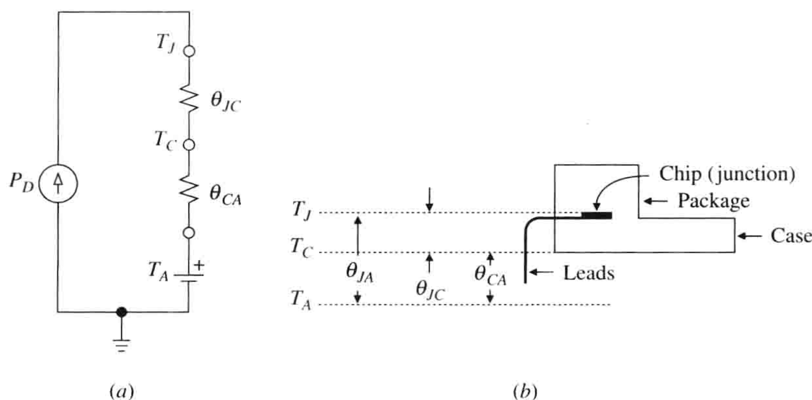
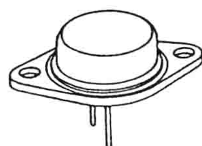


FIGURE 11.24
(*a*) Electrical analog of heat flow. (*b*) Typical package structure operating in free air.



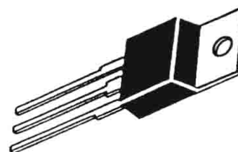


FIGURE 11.25
Two popular power packages. For the μA7800 series, the typical (maximum) thermal-resistance ratings are: TO-3: $\theta_{JC} = 3.5(5.5)\ ^\circ\text{C/W}, \theta_{JA} = 40\ (45)\ ^\circ\text{C/W}$; TO-220: $\theta_{JC} = 3.0\ (5.0)\ ^\circ\text{C/W}$, $\theta_{JA} = 60\ (65)\ ^\circ\text{C/W}$.

EXAMPLE 11.14. (*a*) The data sheets of the 7805 5-V regulator specify $T_{J(\max)} = 150\ ^\circ\text{C}$. Assuming $T_{A(\max)} = 50\ ^\circ\text{C}$, find the maximum power that a TO-220 package operating in free air can dissipate. What is the corresponding case temperature T_C? (*b*) Find the maximum current that can be drawn from the device if $V_I = 8$ V.

Solution.

(*a*) $P_{D(\max)} = (T_{J(\max)} - T_{A(\max)})/\theta_{JA} = (150 - 50)/60 = 1.67$ W. By KVL, $T_C = T_J - \theta_{JC} P_D = 150 - 3 \times 1.67 = 145\ ^\circ\text{C}$.

(*b*) Ignoring the current at the common terminal, we have $P_D \cong (V_I - V_O) I_O$, so $I_O \le 1.67/(8-5) = 0.556$ A.

In the case of free-air operation, heat encounters much more resistance in propagating from case to ambient than from junction to case. The user can reduce θ_{CA} significantly by means of a heatsink. This is a metal structure, usually with fins, that is bonded, clipped, or clamped to the device package to facilitate heat flow from case to ambient. The effect of a heat sink is illustrated in Fig. 11.26. While θ_{JC} remains

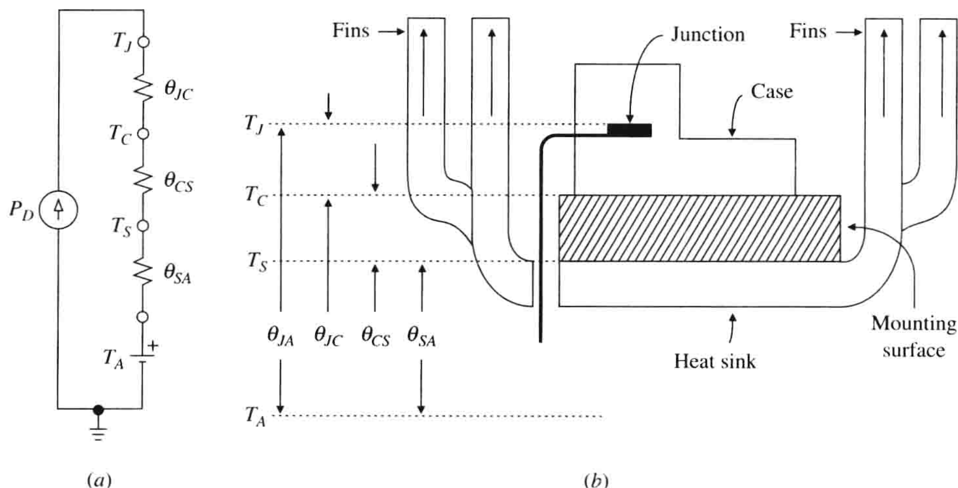


FIGURE 11.26
Heat-flow electrical analog of a package mounted on a heat sink.

the same, θ_{CA} is altered significantly as

$$\theta_{CA} = \theta_{CS} + \theta_{SA} \tag{11.27}$$

where θ_{CS} is the thermal resistance of the mounting surface and θ_{SA} is that of the heatsink. The mounting surface is usually a thin insulating washer of mica or fiberglass to provide electrical isolation between the case, which is internally connected to the collector, and the sink, which is often bonded to the chassis. Usually smeared with heat sink grease to ensure intimate thermal contact, the mounting surface has a typical thermal resistance of less than 1 °C/W.

Heat sinks are available in a variety of shapes and sizes, with thermal resistances ranging from about 30 °C/W for the smaller types to as little as 1 °C/W or less for the truly massive units. Thermal resistance is specified for the case of a heat sink mounted with fins vertical and with unobstructed airflow. Forced air cooling reduces thermal resistance further. In the limiting case of infinite heat sinking and a thermally perfect mounting surface, θ_{CA} would approach zero and the device's ability to expel heat would be limited only by θ_{JC}. The package-heat sink combination best suited to a given application is determined on the basis of the maximum expected power dissipation, the maximum allowable junction temperature, and the maximum anticipated ambient temperature.

EXAMPLE 11.15. A μA7805 regulator is to meet the following requirements: $T_{A(\max)} = 60\,^\circ\text{C}$, $I_{O(\max)} = 0.8$ A, $V_{I(\max)} = 12$ V, and $T_{J(\max)} = 125\,^\circ\text{C}$. Select a suitable package-heat sink combination.

Solution. $\theta_{JA(\max)} = (125 - 60)/[(12 - 5)0.8] = 11.6\,^\circ\text{C/W}$. Use the TO-220 package, which is cheaper and offers better thermal resistance. Then, $\theta_{CA} = \theta_{JA} - \theta_{JC} = 11.6 - 5 = 6.6\,^\circ\text{C/W}$. Allowing 0.6 °C/W for the thermal resistance of the mounting surface, we are left with $\theta_{SA} = 6\,^\circ\text{C/W}$. According to the catalogs, a suitable heat sink example is the IERC HP1 series, whose θ_{SA} rating is in the range of 5 °C/W to 6 °C/W.

Power-Supply Supervisory Circuits

The forms of protection discussed in Section 11.4 safeguard the regulator. A well-designed power-supply system will also include circuitry to safeguard the load and to monitor satisfactory power-supply performance. The functions typically required are *overvoltage* (OV) protection, *undervoltage* (UV) sensing, and *ac line loss* detection. The MC3425 is one of a variety of dedicated circuits known as *power-supply supervisory circuits* designed to assist the designer in this task.

As shown in Fig. 11.27, the circuit consists of a 2.5-V bandgap reference and two comparator channels, one for OV protection and the other for UV detection. The input comparators CMP_1 and CMP_3 have open-collector outputs with 200-μA active pullups. These outputs are externally accessible to allow independent adjustment of the response delays of the two channels in order to prevent false triggering in noisy environments. The delays are established by connecting two capacitors between these outputs and ground, as shown in the subsequent figures.

Under normal conditions these outputs are low. Should, however, an OV or UV condition arise, either CMP_1 or CMP_3 will switch its output BJT off to allow the corresponding delay capacitor to charge by the 200-μA pullup. Once the capacitor voltage reaches V_{REF}, the corresponding output comparator fires, signaling that the emergency condition persisted for the entire delay of that channel. The delay of

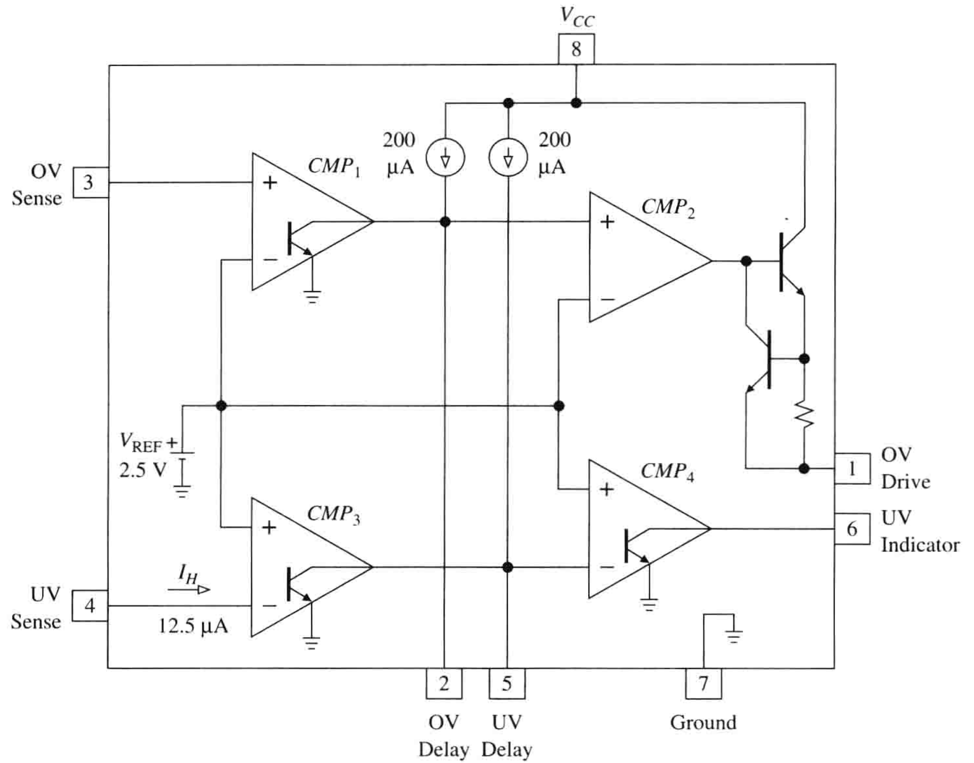


FIGURE 11.27
Simplified diagram of the MC3425 power-supply supervisory/overvoltage-undervoltage protection circuit. (Courtesy of Motorola, Inc.)

either channel is obtained via Eq. (10.2) as $T_{DLY} = C_{DLY}(2.5 \text{ V})/(200 \ \mu\text{A})$, or

$$T_{DLY} = 12{,}500 C_{DLY} \tag{11.28}$$

where C_{DLY} is in farads and T_{DLY} in seconds. For instance, using $C_{DLY} = 0.01 \ \mu$F yields $T_{DLY} = 125 \ \mu$s.

Whereas the UV comparator CMP_4 has an open-collector output, the OV comparator CMP_2 has an overload-protected output booster to drive an external silicon controlled rectifier (SCR) crowbar for emergency power shutdown.

OV/UV Sensing and Line-Loss Detection

Figure 11.28 shows a typical 3425 connection for OV protection and UV sensing. The OV channel trips whenever V_{CC} tries to rise above a level V_{OV} such that $V_{OV}/(1 + R_2/R_1) = V_{REF}$, or

$$V_{OV} = \left(1 + \frac{R_2}{R_1}\right) V_{REF} \tag{11.29}$$

If the OV condition persists for the entire delay T_{OV} as set by C_{OV}, the MC3425 fires the SCR, which in turn shorts out the voltage regulator and blows the fuse, thus

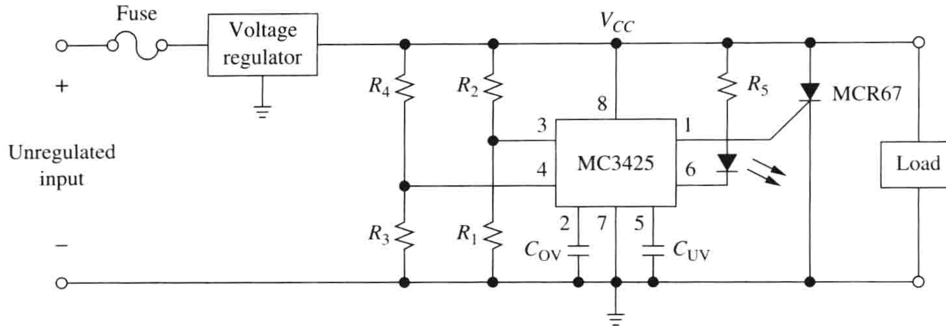


FIGURE 11.28
Overvoltage protection and undervoltage sensing using the MC3425.

protecting the load against prolonged over-voltage and the unregulated input source against prolonged overload.

Likewise, the UV channel trips whenever V_{CC} drops below

$$V_{UV} = \left(1 + \frac{R_4}{R_3}\right) V_{REF} \tag{11.30}$$

Once tripped, CMP_3 also activates an internal circuit that sinks a current $I_H = 12.5\ \mu\text{A}$ from the UV sense input pin. This current is designed to load down the voltage of this pin in order to produce *hysteresis* and, therefore, reduce chatter. The hysteresis width is

$$\Delta V_{UV} = (R_3 \parallel R_4)(12.5\ \mu\text{A}) \tag{11.31}$$

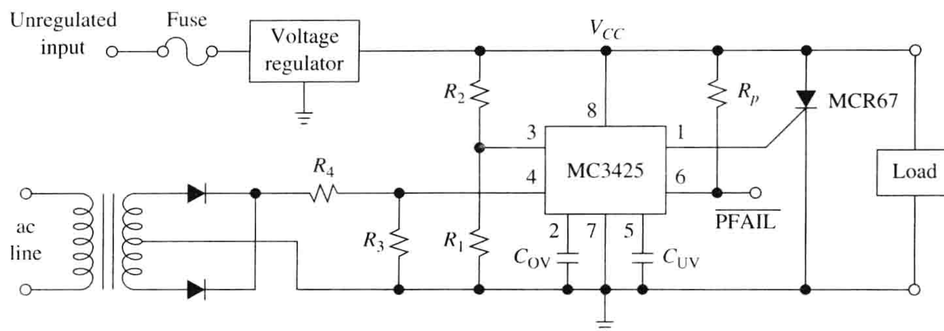
Thus, once CMP_3 fires as a result of V_{CC} dropping below V_{UV}, it remains in that state until V_{CC} rises above $V_{UV} + \Delta V_{UV}$. Unless this happens within the delay T_{UV} as set by C_{UV}, CMP_4 also fires and causes the LED to glow. Once V_{CC} returns above $V_{UV} + \Delta V_{UV}$, CMP_3 returns to the original state and deactivates I_H.

EXAMPLE 11.16. In Fig. 11.28 specify suitable components for an OV trip level of 6.5 V with a 100-μs delay, and a UV trip level of 4.5 V with a 0.25-V hysteresis and a 500-μs delay.

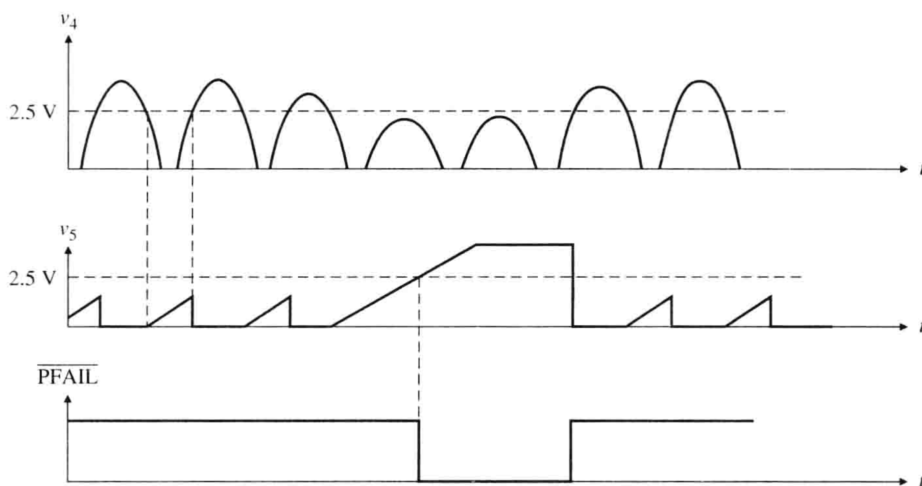
Solution. The above equations give $C_{OV} = 8$ nF, $R_2/R_1 = 1.6$, $R_4/R_3 = 0.8$, $R_3 \parallel R_4 = 20$ kΩ, $C_{UV} = 40$ nF. Use $C_{OV} = 8.2$ nF, $C_{UV} = 43$ nF, $R_1 = 10.0$ kΩ, $R_2 = 16.2$ kΩ, $R_3 = 45.3$ kΩ, $R_4 = 36.5$ kΩ.

In microprocessor-based systems, ac line loss, whether total (*blackout*) or partial (*brownout*), must be detected in time to allow the salvage of vital status information in nonvolatile memory, as well as disable any devices that might be adversely affected by underpowered operation, such as motors and pumps. The circuit of Fig. 11.29*a* monitors the ac line via a center-tapped transformer (which can be the very transformer involved in the generation of the unregulated input to the voltage regulator) and uses the UV channel to detect line loss. Circuit operation is best understood with the help of the waveforms of Fig. 11.29*b*.

The delay capacitor C_{UV} is chosen to be large enough so that, under normal line conditions, it does not have enough time between consecutive ac peaks to charge past 2.5 V. This is also referred to as *retriggerable one-shot operation.* However,



(*a*)



(*b*)

FIGURE 11.29
Overvoltage protection with ac line-loss detection circuit, and typical waveforms.

should the line drop to the extent of causing the peaks at the UV sense pin 4 to drop below the 2.5-V threshold, C_{UV} will fully charge and trigger CMP_4, thus issuing a $\overline{\text{PFAIL}}$ command. This can be used to interrupt the microprocessor and initiate appropriate power-fail routines.

11.6 SWITCHING REGULATORS

As we know, in a linear regulator the series-pass transistor transfers power from V_I to V_O continuously. As depicted in Fig. 11.30*a* for the bipolar case, the BJT operates in the forward-active region, where it acts as a controlled current source dissipating the power $P = V_{CE}I_C + V_{BE}I_B$. Ignoring the base current and the current drawn by the control circuitry compared to the load current I_O, we can write $P \cong (V_I - V_O)I_O$. As already seen, it is precisely this dissipation that limits the efficiency of a linear

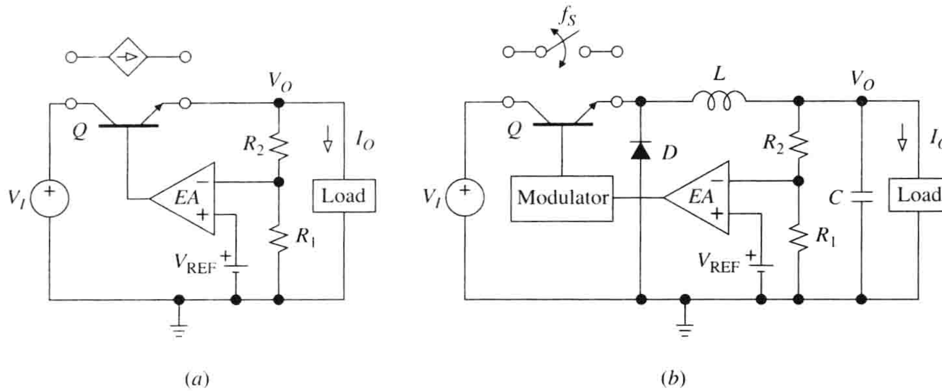


FIGURE 11.30
Linear regulator and switching regulator.

regulator to

$$\eta(\%) = 100\frac{V_O}{V_I}$$

For example, with $V_I = 12$ V and $V_O = 5$ V, we get only $\eta = 41.7\%$.

As we know, proper operation requires that $V_I \geq V_O + V_{DO}$, where V_{DO} is the dropout voltage. A linear regulator of the low-dropout (LDO) type can be made to operate efficiently by powering it with a preregulated voltage near $V_O + V_{DO}$. However, in the absence of any preregulation, V_I may vary well above $V_O + V_{DO}$, making even an LDO regulator inefficient when V_I is at its maximum.

Switching regulators achieve higher efficiency by operating the transistor as a periodically commutated switch as shown in Fig. 11.30*b*. In this case the BJT is either in cutoff, dissipating $P \cong V_{CE}I_C \cong (V_I - V_O) \times 0 = 0$, or in saturation, dissipating $P \cong V_{SAT}I_C$, which is generally small because so is the voltage V_{SAT} across the closed switch. Thus, a switched BJT dissipates much less power than a forward-active BJT. The price for switch-mode operation is the need for a coil to provide a high-frequency transfer of energy packets from V_I to V_O, and a smoothing capacitor to ensure a low output ripple. However, L and C manipulate energy without dissipating any power, at least ideally. Consequently, the combination of switches and low-loss reactive elements makes switching regulators inherently more efficient than their linear counterparts.

Switch-mode regulation is effected by adjusting the duty cycle D of the switch, defined as

$$D = \frac{t_{ON}}{t_{ON} + t_{OFF}} = \frac{t_{ON}}{T_S} = f_S t_{ON} \tag{11.32}$$

where t_{ON} and t_{OFF} are the time intervals during which the transistor is on and off; $T_S = t_{ON} + t_{OFF}$ is the duration of a switch cycle; and $f_S = 1/T_S$ is the operating frequency of the switch. There are two ways of adjusting the duty cycle: (*a*) in *pulse-width modulation* (PWM), f_S is kept fixed and t_{ON} is adjusted; and (*b*) in *pulse-frequency modulation* (PFM), t_{ON} (or t_{OFF}) is fixed and f_S is adjusted. It is apparent that switching regulators require more complex control circuitry than their linear counterparts.

Basic Topologies

If we view the switch-coil-diode combination as a T structure, then, depending on which leg is occupied by the coil, we have the three topologies of Fig. 11.31, called, for reasons to be justified shortly, the *buck, boost,* and *buck-boost* topologies; clearly, the circuit of Fig. 11.30*b* is a buck circuit. Though the topologies are shown for operation with $V_I > 0$, they can readily be configured for $V_I < 0$ by proper reversal of the switch and diode polarities. Moreover, a number of variants[10,11] can be obtained by suitable modification of the coil and switch structures. To gain more insight, we focus on the buck topology, though similar analysis can be applied also to the other topologies.

Assuming $V_I > V_O$ in Fig. 11.31*a*, we can describe buck operation as follows. During t_{ON} the switch closes and connects the coil to V_I. The diode is off, so the situation is as in Fig. 11.32*a*, where V_{SAT} is the voltage drop developed by the closed switch. During this time, current and magnetic energy build up in the coil according to the familiar laws $di_L/dt = v_L/L$ and $w_L = (1/2)Li_L^2$. If V_I and V_O do not change appreciably during a switch cycle, the coil voltage v_L remains relatively

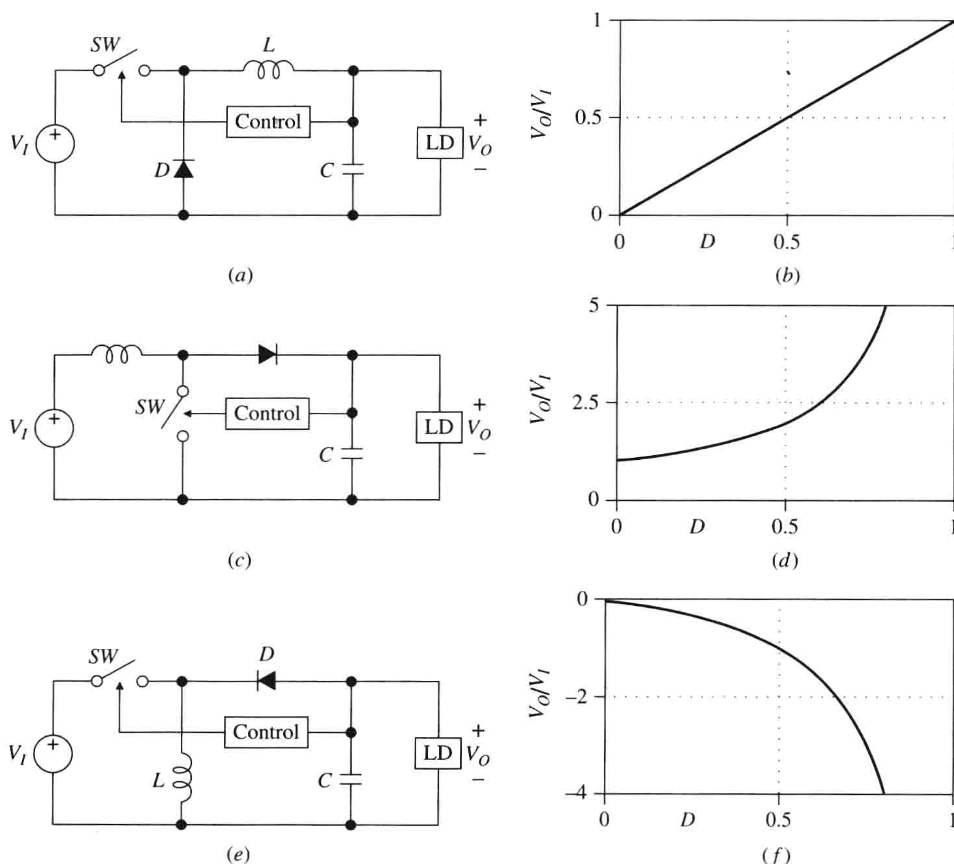


FIGURE 11.31
Basic switching-regulator topologies and idealized V_O/V_I ratios as functions of the duty cycle D: (*a*), (*b*) buck; (*c*), (*d*) boost; (*e*), (*f*) buck-boost.

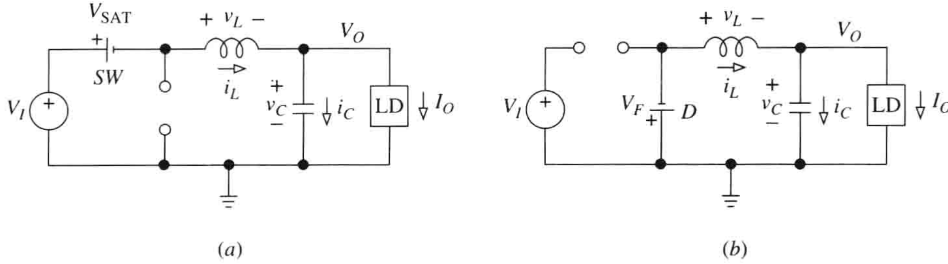


FIGURE 11.32
Equivalent circuits of the buck converter when SW is (a) closed and (b) open.

constant at $v_L = V_I - V_{SAT} - V_O$. We can replace differentials with finite differences and write $\Delta i_L = v_L \Delta t/L$, so during t_{ON} the coil current *increases* by

$$\Delta i_L(t_{ON}) = \frac{V_I - V_{SAT} - V_O}{L} t_{ON} \tag{11.33}$$

Recall from basic physics that current in a coil cannot change instantaneously. Consequently, when the switch is opened, *the coil will develop whatever voltage it takes to maintain the continuity of its current.* As the magnetic field starts to collapse, *di_L/dt changes polarity* and so does v_L, indicating that the coil will swing the voltage of its left terminal negatively until the catch diode turns on to provide a path in which the coil current can continue to flow. The situation is depicted in Fig. 11.32*b*, where V_F is the voltage drop developed by the forward-biased diode. The coil voltage is now $v_L = -V_F - V_O$, indicating a coil current *decrease*

$$\Delta i_L(t_{OFF}) = -\frac{V_F + V_O}{L} t_{OFF} \tag{11.34}$$

Figure 11.33*a* shows the switch, diode, and coil-current waveforms for the case in which the coil current never drops to zero, a situation referred to as *continuous conduction mode* (CCM).

Once the circuit has reached steady-state operation following power turn-on, we have $\Delta i_L(t_{ON}) = -\Delta i_L(t_{OFF}) = \Delta i_L$, where Δi_L is called the *coil current ripple.* Using Eqs. (11.32) through (11.34) gives, for the buck converter,

$$V_O = D(V_I - V_{SAT}) - (1 - D)V_F \tag{11.35}$$

Turning next to the boost topology of Fig. 11.31*c*, we note that the coil voltage, again assumed positive at the left, is $v_L = V_I - V_{SAT}$ during t_{ON}, and $v_L = V_I - (V_F + V_O)$ during t_{OFF}. Proceeding as in the buck case, we find, for the boost converter,

$$V_O = \frac{1}{1 - D}(V_I - DV_{SAT}) - V_F \tag{11.36}$$

Likewise, the coil voltage in Fig. 11.31*e*, assumed positive at the top, is $v_L = V_I - V_{SAT}$ during t_{ON}, and $v_L = V_O - V_F$ during t_{OFF}. Consequently, we have, for the buck-boost converter,

$$V_O = -\frac{D}{1 - D}(V_I - V_{SAT}) + V_F \tag{11.37}$$

In the ideal limits $V_{SAT} \to 0$ and $V_F \to 0$ the above equations simplify, respectively, to the following *lossless* characteristics:

$$V_O = DV_I \qquad V_O = \frac{1}{1 - D}V_I \qquad V_O = -\frac{D}{1 - D}V_I \tag{11.38}$$

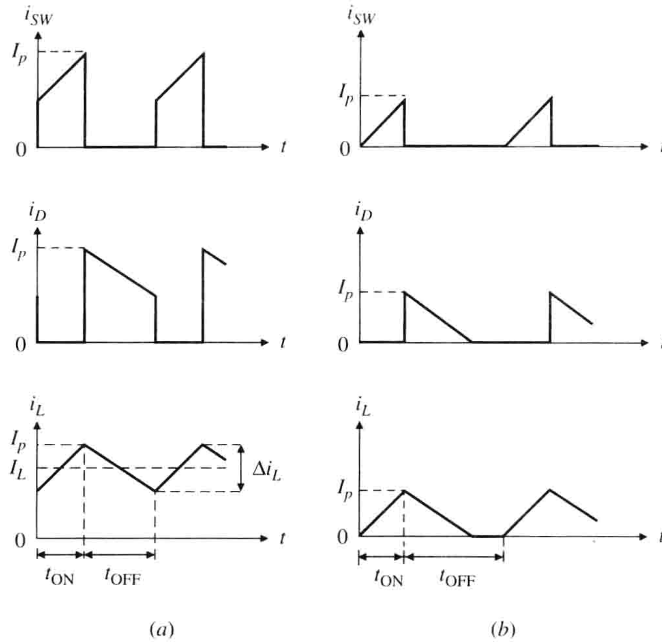


FIGURE 11.33
Current waveforms for the three basic topologies: (*a*) continuous conduction mode (CCM), and (*b*) discontinuous conduction mode (DCM).

These relationships are illustrated pictorially in Fig. 11.31*b*, *d*, and *f*. Considering that $0 < D < 1$, the buck converter yields $V_O < V_I$ and the boost converter $V_O > V_I$, these being the reasons for their names. By analogy with transformers, the buck and boost circuits are also referred to as *step-down* and *step-up* converters. In the buck-boost circuit the output magnitude can be smaller or greater than the input magnitude, depending on whether $D < 0.5$ or $D > 0.5$; moreover, the output polarity is opposite to that of the input, so this converter is also called an *inverting* converter. Note that boost and polarity inversion are not possible with linear regulators!

In the ideal limit of lossless components and zero power dissipation by the control circuitry, a switching regulator would be 100% efficient, giving $P_O = P_I$, or $V_O I_O = V_I I_I$. Writing

$$I_I = (V_O/V_I)I_O \tag{11.39}$$

provides an estimate for the current drawn from the input source.

EXAMPLE 11.17. Given a buck converter with $V_I = 12$ V and $V_O = 5$ V, find D if (*a*) the switch and diode are ideal, and (*b*) $V_{SAT} = 0.5$ V and $V_F = 0.7$ V. (*c*) Repeat (*a*) and (*b*) if 8 V $\le V_I \le$ 16 V.

Solution.

(*a*) By Eq. (11.38), $D = 5/12 = 41.7\%$.
(*b*) By Eq. (11.35), $D = 46.7\%$.
(*c*) The same equations give, for the two cases, $31.2\% \le D \le 62.5\%$, and $35.2\% \le D \le 69.5\%$.

Coil Selection

Two observations should provide better insight into the role of L: (a) The coil must carry some average current $I_L \neq 0$ in order to feed the load; in fact, with reference to the continuous mode shown in Fig. 11.33a, one can prove (see Problem 11.32) that the buck, boost, and buck-boost circuits are characterized, respectively, by

$$I_L = I_O \qquad I_L = \frac{V_O}{V_I} I_O \qquad I_L = \left(1 - \frac{V_O}{V_I}\right) I_O \tag{11.40}$$

(b) In steady state the average coil voltage V_L must be zero.

Should a line or load fluctuation intervene, the controller adjusts the duty cycle D to regulate V_O in accordance with Eq. (11.38), and the coil adjusts I_L to meet the load-current demands in accordance with Eq. (11.40). By the inductance law $i_L = (1/L) \int v_L \, dt$, the coil adjusts its average current I_L by integrating the voltage imbalance brought about by the fluctuation; this adjustment continues until the average coil voltage V_L is driven back to zero.

We can picture the effect of a rise or drop in I_O as an up or down shift of the i_L waveform of Fig. 11.33a. If I_O drops to the point of making $I_L = \Delta i_L/2$, the bottom of the i_L waveform reaches zero. Any further decrease of I_O below this critical value will cause the bottom of the i_L waveform to become clipped, as in Fig. 11.33b, a situation referred to as *discontinuous conduction mode* (DCM). We observe that in CCM V_O depends only on D and V_I, regardless of I_O. By contrast, in DCM V_O depends also on I_O, so D will have to be reduced accordingly by the controller; failing to do so would cause, in the limit of an open-circuited output, $V_O \rightarrow V_I$ for the buck, $V_O \rightarrow \infty$ for the boost, and $V_O \rightarrow -\infty$ for the buck-boost regulators.

To estimate a suitable value of L, it is convenient to assume $V_{SAT} = V_F = 0$. Then, for a buck converter in steady state, Eqs. (11.33) and (11.34) give $t_{ON} = L\Delta i_L/(V_I - V_O)$ and $t_{OFF} = L\Delta i_L/V_O$. Letting $t_{ON} + t_{OFF} = 1/f_S$ gives, for the buck converter,

$$L = \frac{V_O(1 - V_O/V_I)}{f_S \Delta i_L} \tag{11.41}$$

Proceeding in similar manner, we find, for the boost converter,

$$L = \frac{V_I(1 - V_I/V_O)}{f_S \Delta i_L} \tag{11.42}$$

and for the buck-boost converter,

$$L = \frac{V_I/(1 - V_I/V_O)}{f_S \Delta i_L} \tag{11.43}$$

The choice of L is usually a tradeoff between maximum output power with minimum output ripple, and small physical size with fast transient response.[11] Moreover, increasing L for a given I_O will cause the system to go from DCM to CCM. A good starting point is to choose the current ripple Δi_L, and then use the proper equation to estimate L.

There are various criteria for specifying Δi_L. One possibility[11] is to let $\Delta i_L = 0.2 I_{L(\max)}$, where $I_{L(\max)}$ is dictated either by the maximum output current rating of the regulator, as per Eq. (11.40), or by the maximum peak-current rating of the

switch, as per $I_p = I_L + \Delta i_L/2$. The switch rating becomes important especially in step-up situations, where I_L may be considerably larger than I_O. Alternatively, to avoid discontinuous operation, we can let $\Delta i_L = 2I_{O(\min)}$, where $I_{O(\min)}$ is the minimum anticipated load current. Other criteria[11,12] are possible, depending on the type of regulation as well as the objectives of the given application.

Once the value of L has been chosen, a coil must be found that can handle both the peak and rms values of i_L. The peak value is limited by core saturation, for if the coil were to saturate, its inductance would drop abruptly, causing an inordinate rise in i_L during t_{ON}. The rms value is limited by losses in the windings and the core. Though the coil has traditionally been perceived as a very intimidating issue, modern switching-regulator data sheets provide a wealth of useful information to ease coil selection, including coil manufacturers' addresses and specific part numbers.

EXAMPLE 11.18. Specify a coil for a boost converter with $V_I = 5$ V, $V_O = 12$ V, $I_O = 1$ A, and $f_S = 100$ kHz. What is the minimum load current $I_{O(\min)}$ for continuous operation?

Solution. At full load, $I_L = (12/5)1 = 2.4$ A. Let $\Delta i_L = 0.2I_L = 0.48$ A. Then, Eq. (11.44) gives $L = 61\ \mu$H. At full load the coil must withstand $I_p = I_L + \Delta i_L/2 = 2.64$ A, and $I_{rms} = [I_L^2 + (\Delta i_L/\sqrt{12})^2]^{1/2} \cong I_L = 2.4$ A. Moreover, $I_{O(\min)} = 0.1$ A.

Capacitor Selection

To estimate a suitable value of C in the buck topology of Fig. 11.31*a*, we observe that the coil current splits between the capacitor and the load as $i_L = i_C + i_O$. In steady state the average capacitance current is zero and the load current is relatively constant. We can therefore write $\Delta i_C = \Delta i_L$, indicating that the i_C waveform is similar to the i_L waveform, except that i_C is centered about zero, as depicted in Fig. 11.34. The Δi_C ripple causes in turn a Δv_C ripple, which we readily find as $\Delta v_C = \Delta Q_C/C$, where ΔQ_C is the area under the i_C curve over half the clock period $T_S/2$. The area of the triangle is $\Delta Q_C = (1/2) \times (T_S/2) \times (\Delta i_C/2)$. Letting

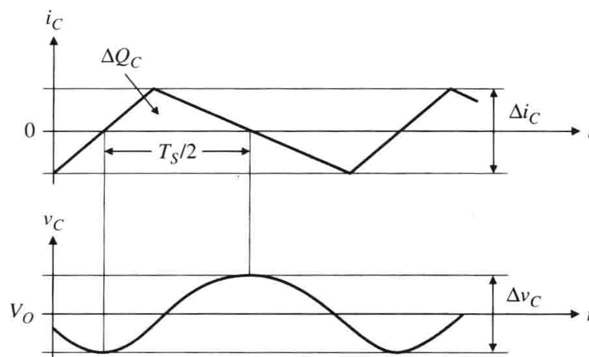


FIGURE 11.34
Capacitor current and voltage waveforms for the buck converter.

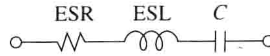


FIGURE 11.35
A practical capacitor has an equivalent series resistance ESR and inductance ESL.

$\Delta i_C = \Delta i_L$ and eliminating ΔQ_C gives, for the buck converter,

$$C = \frac{\Delta i_L}{8 f_S \Delta v_C} \tag{11.44a}$$

In the boost topology of Fig. 11.31*c* the coil is disconnected from the output during t_{ON}, so the load current during this time is supplied by the capacitor. Using Eq. (10.2), we estimate the ripple as $\Delta v_C = I_O t_{ON}/C$. But, $t_{ON} = D/f_S$ and $D = 1 - V_I/V_O$, so we have, for the boost converter,

$$C = \frac{I_O(1 - V_I/V_O)}{f_S \Delta v_C} \tag{11.44b}$$

Similar considerations hold for the buck-boost topology of Fig. 11.31*e*, so

$$C = \frac{I_O(1 - V_I/V_O)}{f_S \Delta v_C} \tag{11.44c}$$

The above equations give C for a specific ripple Δv_C. Practical capacitors exhibit a small *equivalent series resistance* (ESR) and a small *equivalent series inductance* (ESL), as modeled in Fig. 11.35. The ESR contributes an output ripple term of the type $\Delta v_{ESR} = \text{ESR} \times \Delta i_C$, where Δi_C is the capacitor ripple current, indicating the need for low-ESR capacitors. The ripple Δv_C across C in Fig. 11.35 and the ripple Δv_{ESR} across ESR combine to give an overall ripple V_{ro} at the output. For an estimation of the maximum allowed ESR, a reasonable approach[11] is to allow $\frac{1}{3}$ of V_{ro} to come from Δv_C, and $\frac{2}{3}$ of V_{ro} from Δv_{ESR}.

EXAMPLE 11.19. In the boost converter of Example 11.18, specify a capacitor for an output ripple $V_{ro} \cong 100$ mV.

Solution. At full load and with $\Delta v_C \cong (1/3)V_{ro} \cong 33$ mV, Eq. (11.44*b*) gives $C = 177\ \mu$F. For the boost converter we have $\Delta i_C = \Delta i_D = I_p$, so at full load $\Delta i_C = 2.64$ A. Then, ESR $= (67\text{ mV})/(2.64\text{ A}) \cong 25\text{ m}\Omega$.

The C and ESR requirements may be difficult to meet simultaneously, so we can either increase the size of the capacitor, since larger capacitors tend to have smaller ESRs, or we can filter out the existing ripple with an additional LC stage at the output.

A well-constructed switching regulator will include an LC filter also at the input, both to ease the output-impedance requirements of the source V_I and to prevent the injection of electromagnetic interference (EMI) upstream of the regulator. We observe that the most taxing situation for a capacitor is when it is in series with either the switch or the diode. When it is in series with the coil, as at the input of the boost or at the output of the buck topology, the filtering action provided by the coil itself results in a smoother waveform. It follows that the buck converter enjoys the lowest output ripple of the three topologies.

Efficiency

The efficiency of a switching regulator is found as

$$\eta(\%) = \frac{P_O}{P_O + P_{\text{diss}}} \tag{11.45}$$

where $P_O = V_O I_O$ is the power delivered to the load, and

$$P_{\text{diss}} = P_{SW} + P_D + P_{\text{coil}} + P_{\text{cap}} + P_{\text{controller}} \tag{11.46}$$

is the sum of the losses in the switch, the diode, the coil, the capacitor, and the switch controller.

Switch loss is the sum of a *conduction* component and a *switching* component, or $P_{SW} = V_{\text{SAT}} I_{SW} + f_S W_{SW}$. The conduction component is due to the nonzero voltage drop V_{SAT}; for the case of a saturating BJT switch this component is found as $V_{CE(\text{sat})} I_{SW(\text{avg})}$, and for the case of a FET switch as $r_{DS(\text{on})} I^2_{SW(\text{rms})}$. The switching component is due to the nonzero rise and fall times of the voltage and current waveforms of the switch; the resulting waveform overlap causes the per-cycle dissipation of an energy packet[12] $W_{SW} \cong 2\Delta v_{SW} \Delta i_{SW} t_{SW}$, where Δv_{SW} and Δi_{SW} are the switch voltage and current changes, and t_{SW} is the effective overlap time.

Diode dissipation is likewise[11] $P_D = V_F I_{F(\text{avg})} + f_S W_D$, $W_D \cong V_R I_F t_{\text{RR}}$, where V_R is the diode reverse voltage, I_F the forward current at turn-off, and t_{RR} the reverse recovery time. Schottky diodes are good choices because of their inherently lower voltage drop V_F and the absence of charge-storage effects.

Capacitor loss is $P_{\text{cap}} = \text{ESR} I^2_{C(\text{rms})}$. Coil loss consists of two terms, namely, the copper loss $R_{\text{coil}} I^2_{L(\text{rms})}$ in the coil resistance, and core losses, which depend on the coil current as well as f_S. Finally, the controller contributes $V_I I_Q$ where I_Q is the average current it draws from V_I, exclusive of the switch.

EXAMPLE 11.20. A buck converter with $V_I = 15$ V, $V_O = 5$ V, $I_O = 3$ A, $f_S = 50$ kHz, and $I_Q = 10$ mA, uses a switch with $V_{\text{SAT}} = 1$ V and $t_{SW} = 100$ ns, a diode with $V_F = 0.7$ V and $t_{\text{RR}} = 100$ ns, a coil with $R_{\text{coil}} = 50$ mΩ and $\Delta i_L = 0.6$ A, and a capacitor with ESR = 100 mΩ. Assuming core losses of 0.25 W, find η and compare with a linear regulator.

Solution. Eq. (11.35) gives $D = 38.8\%$. Then, $P_{SW} \cong V_{\text{SAT}} D I_O + 2 f_S V_I I_O t_{SW} = 1.16 + 0.45 = 1.61$ W; $P_D \cong V_F(1 - D) I_O + f_S V_I I_O t_{\text{RR}} = 1.29 + 0.22 = 1.51$ W; $P_{\text{cap}} = \text{ESR}(\Delta i_L/\sqrt{12})^2 = 3$ mW; $P_{\text{coil}} = R_{\text{coil}} \times (\Delta i_L/\sqrt{12})^2 + 0.25 \text{ W} \cong 0.25$ W; $P_{\text{controller}} = 15 \times 10 = 0.15$ W; $P_O = 5 \times 3 = 15$ W; $P_{\text{diss}} = 3.52$ W; $\eta = 81\%$.

A linear regulator would have $\eta = 5/15 = 33\%$, indicating that to deliver 15 W of useful power it would dissipate 30 W, while the switching regulator of our example dissipates only 3.52 W.

11.7 THE ERROR AMPLIFIER

As we know, the primary function of the error amplifier (EA) in a switching regulator is to accept a scaled version βV_O of the output voltage, compare it against an internal reference voltage V_{REF}, and issue to the switch modulator whatever control voltage

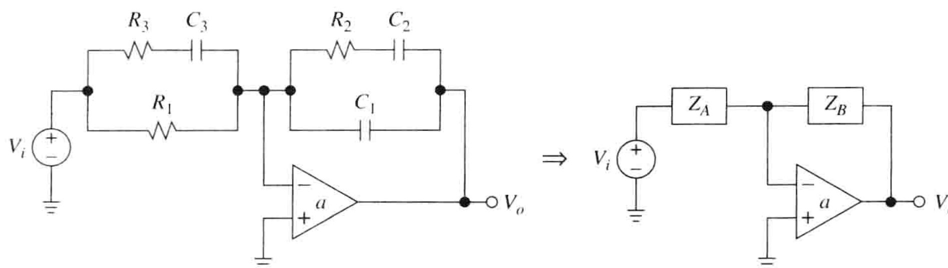


FIGURE 11.36
AC equivalent of the Type 3 error amplifier (EA).

v_{EA} it takes to make βV_O track V_{REF}, or, equivalently, to make $V_O = V_{\text{REF}}/\beta$. If there were no other delays around the control loop, a plain integrator would do, provided we set its unity-gain frequency f_0 well below the switching frequency f_S (say, $f_0 \cong f_S/10$) to help filter out the switching noise. So, with a phase shift of $-180°$ due to negative feedback, and an additional shift of $-90°$ due to the integrator, we would have a phase margin of $360 - 180 - 90 = 90°$. But the loop of a switcher includes also a capacitor and an inductor, whose tendency is to introduce additional delays and thus erode the phase margin. This requires that the integrator's response in the vicinity of the crossover frequency be suitably *shaped* in order to ensure an acceptable *noise margin* for the overall loop (for this reason, the EA is also called a *compensator*).

EAs are categorized as Type 1, Type 2, and Type 3. As we will see, Type 1 is a particular case of Type 2, which in turn is a particular case of Type 3, so let us discuss Type 3, the most general type. Its popular embodiment of Fig. 11.36 uses an op amp with an adequate GBP to provide the low-frequency gain necessary to make $V_O = V_{\text{REF}}/\beta$, and the impedance pair Z_A-Z_B to establish the frequency profile necessary for an adequate phase margin. Though the transfer function can be derived mathematically as

$$H_{\text{EA}}(j\omega) = -\frac{Z_B(j\omega)}{Z_A(j\omega)} = -\frac{(1 + j\omega/\omega_{z1})(1 + j\omega/\omega_{z2})}{(j\omega/\omega_0)(1 + j\omega/\omega_{p1})(1 + j\omega/\omega_{p2})} \qquad (11.47)$$

(see Problem 11.37), we can gain far better insight if we examine the individual impedances using physical inspection. This task is facilitated further by the fact that EAs are implemented under the conditions

$$C_2 \gg C_1 \qquad R_1 \gg R_3 \qquad (11.48)$$

Consider first Z_B. At low frequencies, where the impedances of C_1 and C_2 have much greater magnitude than R_2, we can ignore R_2 and write $|Z_B| \to 1/[\omega(C_1 + C_2)] \cong 1/(\omega C_2)$, by Eq. (11.48). At high frequencies, where the capacitive impedances are much smaller than R_2, we can ignore C_2 compared to R_2, and in turn ignore R_2 compared to C_1, so now $|Z_B| \to 1/(\omega C_1)$. The transition from one asymptote to the other must occur over a region where Z_B is flat. In fact, this is the intermediate-frequency region where $Z_B = R_2$. As depicted at the top of Fig. 11.37, Z_B exhibits a zero-pole frequency pair. These are the frequencies at which $1/(\omega_{z1} C_2) = R_2$ and $1/(\omega_{p2} C_1) = R_2$.

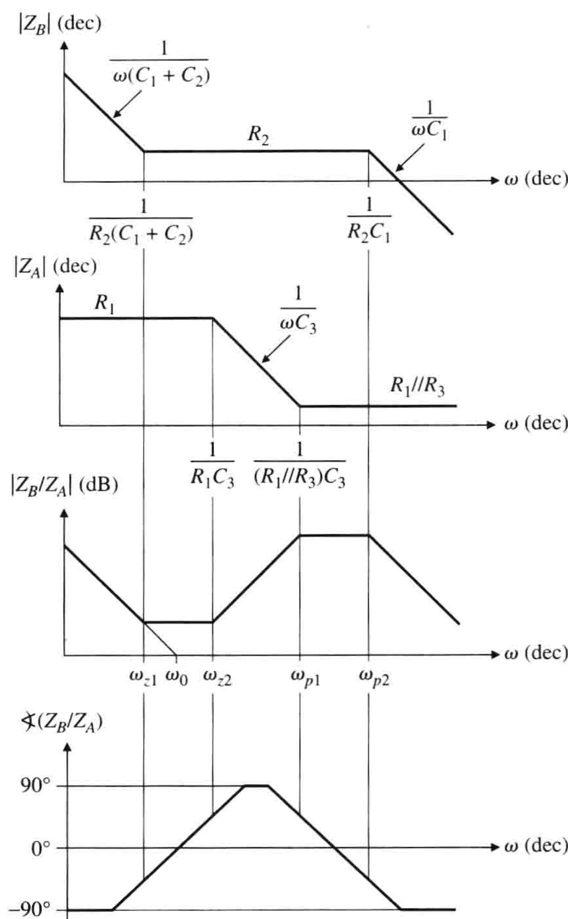


FIGURE 11.37
Using the linearized Bode plots of the individual impendaces of Fig. 11.36 to construct the magnitude and phase plots of their ratio.

Applying similar reasoning to Z_A, we can state that at low frequencies we have $Z_A \rightarrow R_1$ and at high frequencies we have $Z_A \rightarrow R_1 // R_3 \cong R_3$, by Eq. (11.48). The transition from one asymptote to the other must occur over a region of decreasing Z_A with ω. In fact, this is the intermediate-frequency region where $|Z_A| = 1/(\omega C_3)$. As seen in Fig. 11.37, Z_A exhibits a pole-zero frequency pair, which in turn is a zero-pole pair for $1/Z_A$. These are the frequencies at which $1/(\omega_{z2} C_3) = R_1$ and $1/(\omega_{p1} C_3) = R_3$.

Having the logarithmic plots of $|Z_A|$ and $|Z_B|$ in hand, we construct the logarithmic plot of their ratio $|Z_B/Z_A|$ as the *difference* between the individual logarithmic values, multiplied by 20 to convert to decibels. Finally, using the phase-slope correspondence of Eq. (8.9), we construct the phase plot shown at the bottom. It is apparent that the EA exhibits integrator behavior both at low frequencies (where it is designed to provide the high gain necessary to make $V_O = V_{\mathrm{REF}}/\beta$), and at high

frequencies (where it is designed to provide the low-pass function needed to filter out switching noise). At intermediate frequencies the response has been reshaped through the creation of two zero frequencies, whose cumulative effect is to raise phase from $-90°$ to $+90°$! We can certainly use this phase lead to neutralize the phase lag of two additional loop poles, or of an additional pole and a right-half-plane zero, as we shall see below. In light of the above considerations, the characteristic frequencies of the EA are

$$\omega_{z1} \cong \frac{1}{R_2C_2} \qquad \omega_{z2} \cong \frac{1}{R_1C_3} \qquad \omega_{p1} \cong \frac{1}{R_3C_3} \qquad \omega_{p2} \cong \frac{1}{R_2C_1} \tag{11.49a}$$

Also relevant is the extrapolated unity-gain frequency ω_0 of the low-frequency integrator segment,

$$\omega_0 = \frac{1}{R_1(C_1 + C_2)} \cong \frac{1}{R_1C_2} \tag{11.49b}$$

which is used as a scaling factor to shift the magnitude plot up or down without affecting phase.

As mentioned, the Type 3 EA is the most general of the three types. Type 2 is obtained from Type 3 simply by discarding R_3 and C_3 so that $Z_A = R_1$. In this case the profile of $|Z_B/Z_A|$ is similar to that of $|Z_B|$, indicating a phase boost of only $+90°$. Type 2 is used when the feedback loop includes only one additional pole frequency, as in the case of peak current mode control for buck converters, to be investigated in the next section. Finally, Type 1 EA is obtained from Type 3 by discarding all components except for R_1 and C_1, thus reducing the EA to a plain integrator. This type finds use in linear regulators. It ought to be pointed out that the implementation of Fig. 11.38 is not unique (see Problem 11.38 for an alternative implementation often proposed in the application notes). Also, the op amp used to implement the EA must have an adequate GBP to avoid degrading the phase margin too much (see Problem 11.41).

11.8 VOLTAGE MODE CONTROL

Voltage mode control (VMC), exemplified in Fig. 11.38 for the buck topology, synthesizes the switch drive by modulating a sawtooth wave v_{ST} of frequency f_S with the voltage v_{EA} produced by the error amplifier (modulation is achieved by means of the voltage comparator *CMP* along the lines already discussed in connection with Fig. 9.18). The modulation waveforms are shown in Fig. 11.39. Note that in the present converter rendition, typical of today's CMOS technology, the catch diode has been replaced by M_n, a low-channel resistance MOSFET designed to ensure an even lower voltage drop than a Schottky diode. This is to reduce power losses and thus improve efficiency. Note also that the catch FET M_n must be synchronized in push-pull fashion with the switch proper M_p (hence, the name), and that it must be driven in break-before-make fashion to avoid establishing a low-resistance path from the source V_I to ground (these functions are performed by the break-before-make driver circuit, whose details are beyond our scope here). Note also the generalized load model, consisting of a current sink I_{LD} in parallel with a resistance R_{LD}.

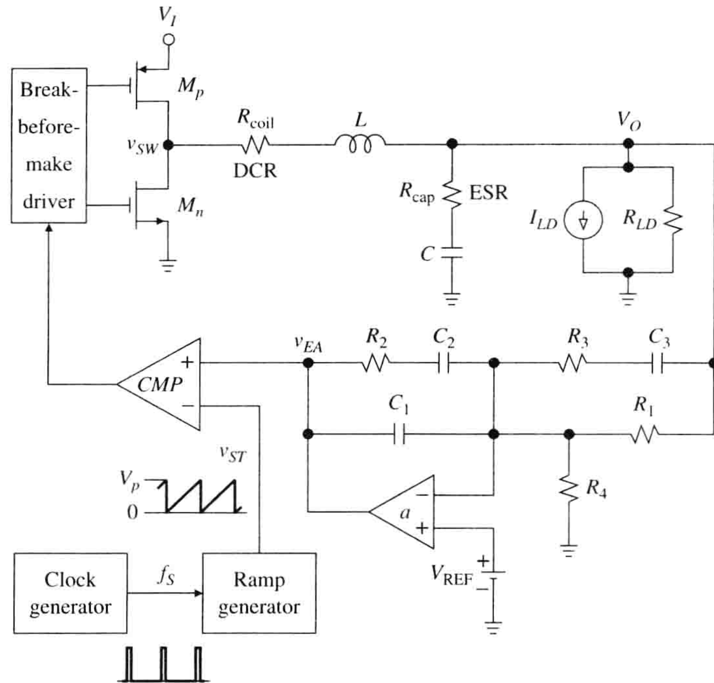


FIGURE 11.38
Synchronous buck converter using voltage mode control (VMC).

We now wish to investigate the stability of the circuit and design the EA for an adequate phase margin. Central to this task is the loop gain T, which, by definition, we find by going around the loop as

$$T = -\frac{V_{sw}}{V_{ea}} \times \frac{V_o}{V_{sw}} \times \frac{V_{ea}}{V_o} \tag{11.50}$$

where V_{sw}, V_{ea}, and V_o are the Laplace transforms of v_{SW}, v_{EA}, and v_O. In the absence of the inductor, the switch would produce the open-circuit, rail-to-rail square wave $v_{SW(oc)} = DV_I$ with the duty cycle $D = v_{EA}/V_p$, or $v_{SW(oc)} = (V_I/V_p)v_{EA}$. Consequently, to simplify ac analysis, we apply Thévenin's theorem

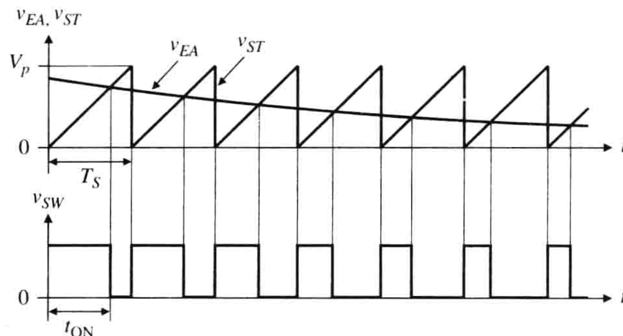


FIGURE 11.39
Voltage modulation in the buck converter of Fig. 11.38.

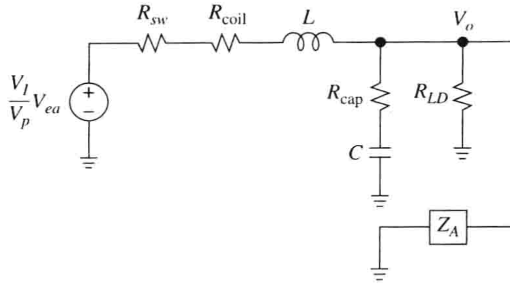


FIGURE 11.40
AC equivalent of coil-capacitor structure of Fig. 11.38.

and use the equivalent of Fig. 11.40, where R_{sw} is the effective resistance presented by the CMOS switch. (Note that the dc sink I_{LD} does not appear in this ac equivalent, and that Z_A is terminated on the virtual ac ground presented by the inverting input of the op amp.) It is left as an exercise (see Problem 11.39) to prove that, so long as loading by R_{LD} and Z_A is negligible, we have

$$T(j\omega) = \frac{V_I}{V_p} \times \frac{1 + j\omega/\omega_{ESR}}{1 - (\omega/\omega_{LC})^2 + (j\omega/\omega_{LC})/Q} \times \frac{(1 + j\omega/\omega_{z1})(1 + j\omega/\omega_{z2})}{(j\omega/\omega_0)(1 + j\omega/\omega_{p1})(1 + j\omega/\omega_{p2})} \qquad (11.51)$$

where

$$\omega_{LC} = \frac{1}{\sqrt{LC}} \qquad \omega_{ESR} = \frac{1}{R_{cap}C} \qquad Q = \frac{\sqrt{L/C}}{R_{sw} + R_{coil} + R_{cap}} \qquad (11.52)$$

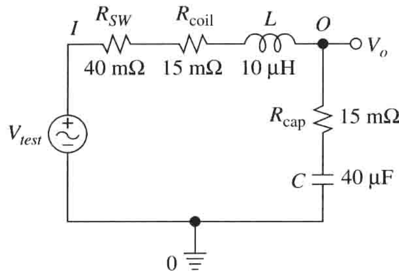
Note that for $\omega \ll \omega_{ESR}$, the LC structure exhibits the familiar second-order low-pass response H_{LP}. However, at high frequencies, where C acts as a short circuit compared to its own ESR, the structure turns into a first-order LR circuit. The borderline between the two cases is the frequency at which $|Z_C(j\omega_{ESR})| = R_{coil}$. Aptly referred to as left-half-plane zero (LHPZ) frequency, ω_{ESR} marks the point where slope changes from -40 dB/dec to -20 dB/dec, and phase is on its way to get boosted by $+90°$.

EXAMPLE 11.21. (*a*) Assuming $V_I = 10$ V, $V_p = 1$ V, $V_{REF} = 1.282$ V, and $f_S = 500$ kHz for the buck converter of Fig. 11.38, specify suitable component values for $V_O = 3.3$ V and $I_O = 2.5$ A, under the constraint of an output ripple not exceeding 10 mV. (*b*) Assuming $R_{sw} = 40$ mΩ, calculate ω_{LC}, Q, and ω_{ESR}. (*c*) Use PSpice to display the Bode plots (magnitude and phase) of the transfer function from the source at the left of Fig. 11.40 to the output node at the right.

Solution.

(*a*) Imposing $3.3 = (1 + R_1/R_4)1.282$ gives $R_1/R_4 = 1.57$. Use $R_1 = 10.0$ kΩ and $R_4 = 6.35$ kΩ. Imposing $\Delta i_L = 0.2I_O = 0.2 \times 2.5 = 0.5$ A and using Eq. (11.41), we get

$$L = \frac{3.3(1 - 3.3/10)}{500 \times 10^3 \times 0.5} = 8.84\ \mu\text{H}$$



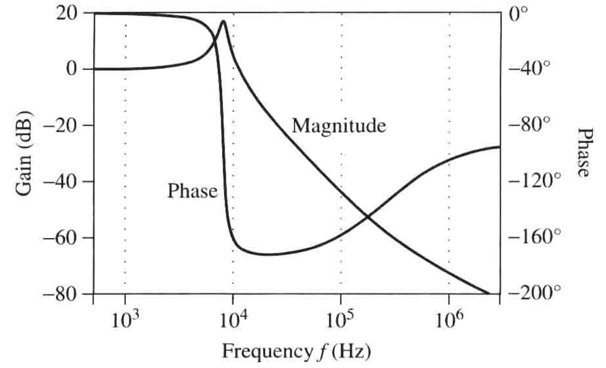


FIGURE 11.41
(*a*) PSpice circuit to generate (*b*) the Bode plots of the coil-capacitor filter.

Use $L = 10\ \mu$H, rated to withstand an average current of $I_L = 2.5$ A and a peak current of $I_p = I_L + \Delta i_L/2 = 2.5 + 0.5/2 = 2.75$ A. Use a 10-μH coil having a DCR of 15 mΩ. Assuming (1/3) of the output ripple $\Delta v_O (\leq 10$ mV) is due to C, use Eq. (11.44*a*) to calculate

$$C = \frac{0.5}{8 \times 500 \times 10^3 \times (1/3)10 \times 10^{-3}} = 37.5\ \mu\text{F}$$

Use $C = 40\ \mu$F. Assuming the remainder (2/3) of $\Delta v_O (= 6.7$ mV) is due to the ESR, estimate $R_{\text{cap}} = (6.7\text{ mV})/(0.5\text{ A}) = 13.3$ mΩ. Use a 40-μF capacitor having an ESR of 15 mΩ.

(*b*) Using Eq. (11.52), we get

$$f_{LC} = \frac{1/(2\pi)}{\sqrt{10 \times 40 \times 10^{-6}}} = 8.0\text{ kHz} \quad f_{\text{ESR}} = \frac{1/(2\pi)}{15 \times 10^{-3} \times 40 \times 10^{-6}} = 265\text{ kHz}$$

$$Q = \frac{\sqrt{10/40}}{(40 + 15 + 15)10^{-3}} = 7.14 (= 17\text{ dB})$$

(*c*) Using the PSpice circuit of Fig. 11.41*a*, we get the plots of Fig. 11.41*b*, showing the effect of the LHPZ.

Error Amplifier Design

As seen above, the LC structure introduces a pole pair inside the loop, along with a zero due to the capacitor's ESR. The pole pair yields a phase shift of $-180°$, and the zero a phase shift of $+90°$. Depending on the location of the crossover frequency relative to the above LC pole pair and the ESR zero, we may need a Type 3 or, at best, a Type 2 compensator. The strategy for placing the compensator's characteristic frequencies is not unique (in fact, different application notes provide somewhat differing rules that may easily confuse the neophyte). Nevertheless, we can say that (*a*) the crossover frequency ω_x is placed typically a decade or so below the switching frequency ω_S, or $\omega_x \cong \omega_S/10$, (*b*) the zero frequencies ω_{z1} and ω_{z2} are placed in the vicinity of ω_{LC} to make their combined phase lead combat the phase lag due to the LC pole pair, and (*c*) the pole frequencies ω_{p1} and ω_{p2} are placed near

ω_S. If ω_{ESR} happens to be sufficiently close to ω_{LC}, then ω_{p1} can be placed right at ω_{ESR} to get a pole-zero cancellation so we can use just a Type 2 compensator.

EXAMPLE 11.22. (*a*) Design an EA for the regulator of Example 11.21 under the following constraints: $f_x = f_S/10$, $f_{z1} = f_{z2} = f_{LC}$, and $f_{p1} = f_{p2} = 4f_x$. (*b*) Use PSpice to generate the Bode plots of the loop gain and measure the phase margin. (*c*) Use PSpice to show the regulator response to a 1-A load-current step.

Solution.

(*a*) We have $f_x = f_S/10 = 500/10 = 50$ kHz, $f_{z1} = f_{z2} = f_{LC} = 8$ kHz, $f_{p1} = f_{p2} = 4f_x = 200$ kHz. To find the required f_0, impose $|T(jf_x)| = 1$ in Eq. (11.51),

$$1 = \frac{10}{1} \times \sqrt{\frac{1 + (50/265)^2}{\left[1 - (50/8)^2\right]^2 + [(50/8)/7.14]^2}} \times \frac{1}{50/f_0} \times \frac{1 + (50/8)^2}{1 + (50/200)^2}$$

This gives $f_0 = 5.0$ kHz. Starting out with $R_1 = 10$ kΩ, turn around Eqs. (11.49*a*) and (11.49*b*) and calculate

$$C_2 = \frac{1}{\omega_0 R_1} = \frac{1}{2\pi 5 \times 10^3 \times 10^4} = 3.2 \text{ nF}$$

$$R_2 = \frac{1}{\omega_{z1} C_2} = \frac{1}{2\pi 8 \times 10^3 \times 3.2 \times 10^{-9}} = 6.2 \text{ k}\Omega$$

$$C_3 = \frac{1}{\omega_{z2} R_1} = \frac{1}{2\pi 8 \times 10^3 \times 10^4} = 2.0 \text{ nF}$$

$$R_3 = \frac{1}{\omega_{p1} C_3} = \frac{1}{2\pi 200 \times 10^3 \times 2 \times 10^{-9}} = 0.4 \text{ k}\Omega$$

$$C_1 = \frac{1}{\omega_{p2} R_2} = \frac{1}{2\pi 200 \times 10^3 \times 6.2 \times 10^3} = 128 \text{ pF}$$

(*b*) Whether simulating the circuit via PSpice or testing it in the lab, use the injection technique of Section 8.2 to measure the loop gain. The PSpice circuit of Fig. 11.42 uses the VCVS EOA to simulate an op amp with a gain of 80dB, and the VCVS Emod to simulate the modulator gain V_I/V_p = 10 V/V. Since this source presents infinite input resistance, breaking the loop immediately upstream of it will require just one voltage-type injection. The loop gain is then $T = -V_r/V_f$. The result of the simulation, shown in Fig. 11.43*a*, reveals a crossover frequency of $f_x = 49.8$ kHz, where $\measuredangle T = -123.4°$, so the phase margin is $\phi_m = 180 - 123.4 = 56.6°$.

(*c*) For the step response we use the same circuit, but with a step current out of the output node. The result, shown in Fig. 11.43*b*, consists of an initial abrupt drop of 15 mV due to the 1-A current flowing initially out of R_{cap}. Afterward, the loop takes over, causing a bit of overshoot, as shown. Regardless of the root placement chosen for the EA, the calculated component values should be taken only as a starting point. A careful engineer will observe the step response experimentally, and fine-tune the EA in order to optimize the transient response for the particular application at hand.

Before concluding, we wish to point out that the above discussion assumes V_I to be a relatively constant voltage for the buck regulator to scale down to the desired value of V_O. Should V_I be widely variable, the ratio V_I/V_p will also vary, causing $|T|$ to move up or down. This, in turn, will move f_x up or down the frequency spectrum, upsetting the phase margin of the circuit. Modern regulators avoid this drawback by making V_p track V_I and thus maintain a stable V_I/V_p ratio.

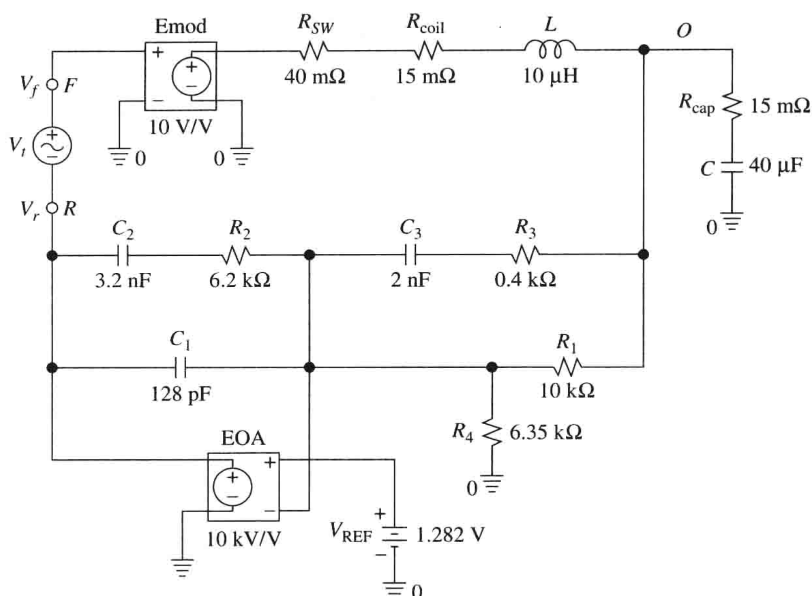


FIGURE 11.42
PSpice circuit to plot the loop gain of the buck regulator of Example 11.22.

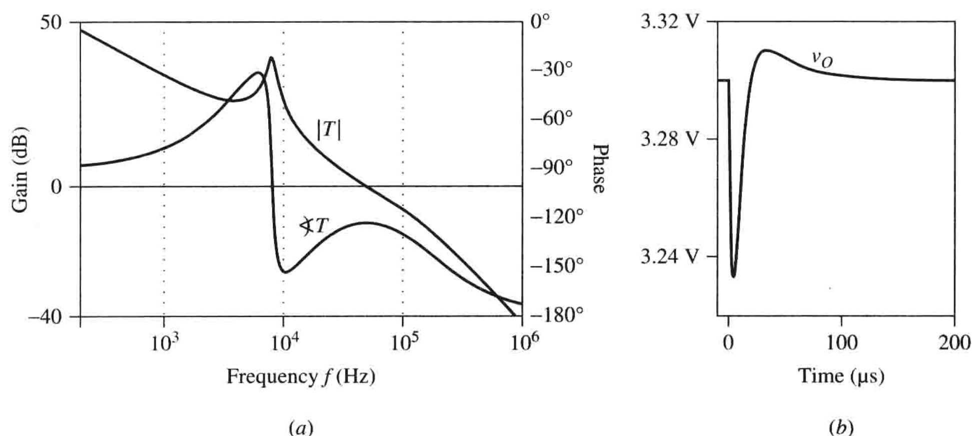


FIGURE 11.43
(*a*) Loop gain for the buck regulator of Fig. 11.42, and (*b*) its response to a 1-A output current step.

11.9 PEAK CURRENT MODE CONTROL

The aim of current mode control (CMC) is to drive the inductor on a cycle-by-cycle basis so as to force it to act as a voltage-controlled current source (VCCS). Two important advantages accrue from this form of control: (*a*) by intervening on the

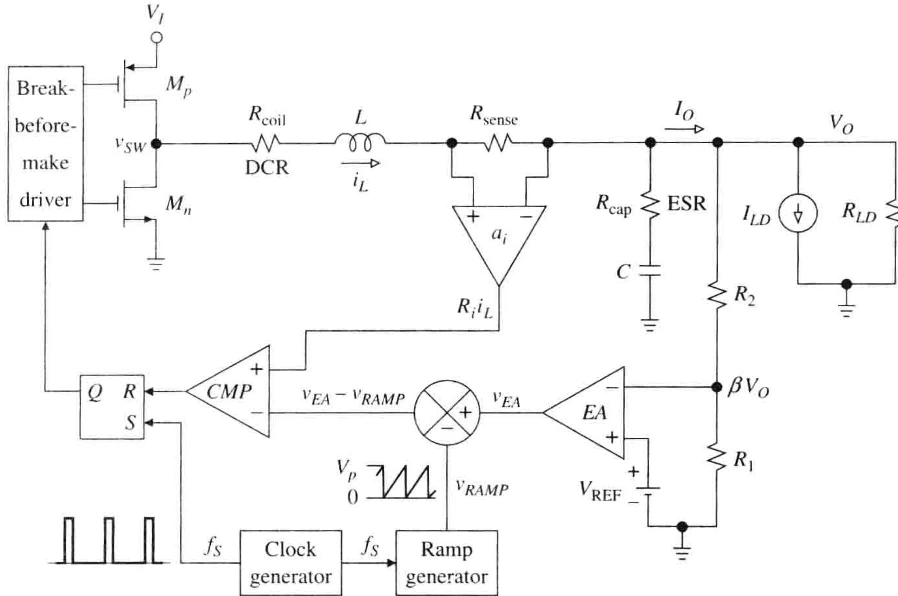


FIGURE 11.44
Buck converter with peak current mode control (PCMC).

inductor current directly, without waiting for the loop delay of voltage mode control (VMC), we are effectively removing the inductor delay from the loop, thus making loop compensation easier; (*b*) direct cycle-by-cycle current control makes it easier to protect the inductor against excessive current buildup in case of malfunction.

PCMC for the Buck Converter

In order to control i_L we need to sense it. In the buck converter of Fig. 11.44, i_L is sensed via a small (few tens of mΩ) series resistor R_{sense}, whose voltage drop is then magnified by an amplifier with gain a_i to convert i_L to the voltage $R_i i_L$, where

$$R_i = a_i \times R_{sense} \tag{11.53}$$

is the overall gain of the current-to-voltage conversion, in V/A, or ohms. This voltage is then fed to the noninverting input of the voltage comparator *CMP*. (Other forms of current sensing are in use, such as utilizing the channel resistance of M_p as the sensing resistor, or by inferring the voltage drop across R_{coil} via a suitable network across the entire coil.) Current sensing comes with a notorious drawback, namely, the fact that R_{sense} must of necessity be kept small to limit power dissipation for the sake of efficiency makes current sensing much more sensitive to noise than voltage mode control, where the duty cycle is created via a much heftier sawtooth wave.

The most common form of current control is the *peak current mode control* (PCMC) depicted in Fig. 11.45. A cycle initiates when a clock pulse of frequency f_S and period $T_S = 1/f_S$ sets the flip-flop. This turns on M_p, causing i_L to ramp up with a slope of

$$S_n = \frac{V_I - V_O}{L} \tag{11.54a}$$

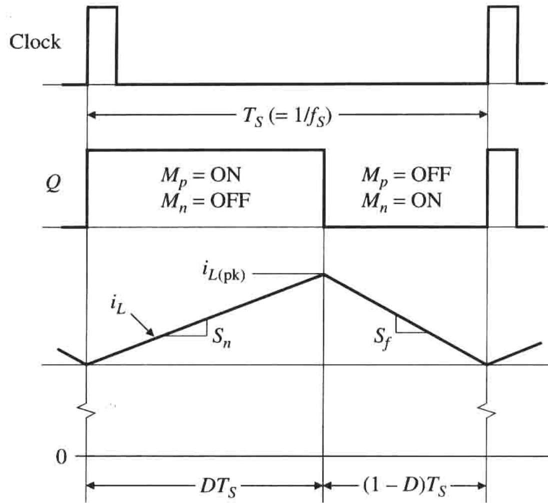


FIGURE 11.45
Timing of PCM control for the buck converter.

During this time, the energy $w_L = (1/2)Li_L^2$ builds up in the inductor. Once i_L reaches a peak value $i_{L(pk)}$ established by the error amplifier, the voltage $R_i i_L$ catches up with v_{EA}, causing the comparator to fire and reset the flip-flop. This turns off M_p while turning on M_n, in a break-before-make fashion to avoid simultaneous conduction. At this instant the magnetic field inside the inductor starts to collapse, so for the rest of the cycle i_L will ramp down with a slope of

$$S_f = \frac{-V_O}{L} \tag{11.54b}$$

By the end of the cycle an energy packet will have been transferred from the inductor to the capacitor-load combination.

As the capacitor continues to receive energy packets, its voltage will build up until, after a sufficient number of cycles, an equilibrium condition is reached, aptly referred to as *steady state,* whereby the average capacitor voltage V_C will no longer increase (though its instantaneous value v_C will continue to alternate around this average value). Consequently, the capacitor current must average to zero, a condition referred to as the *capacitor amp-second balance*. With $I_C = 0$, we have, by KCL, $I_O = I_L$. In steady state the inductor satisfies the dual condition known as the *inductor volt-second balance*, whereby its voltage averages to zero (though its instantaneous value v_L will continue to alternate around this average value).

Should a change in the line voltage V_I intervene, the buck converter reacts by adjusting the duty cycle D so as to keep V_O at the prescribed value. Likewise, should a change in the load current I_O intervene, the converter adjusts D so as to move the average inductor current I_L up or down until the condition $I_L = I_O$ is met. Clearly, we expect the converter to exhibit a good line and load regulation.

We wish to visualize a cycle in terms of the EA output v_{EA}. First, assume $v_{RAMP} = 0$ in Fig. 11.44 (this is referred to as operation without slope compensation, to be discussed shortly), so that *CMP* fires when $R_i i_L$ catches up with v_{EA}, or,

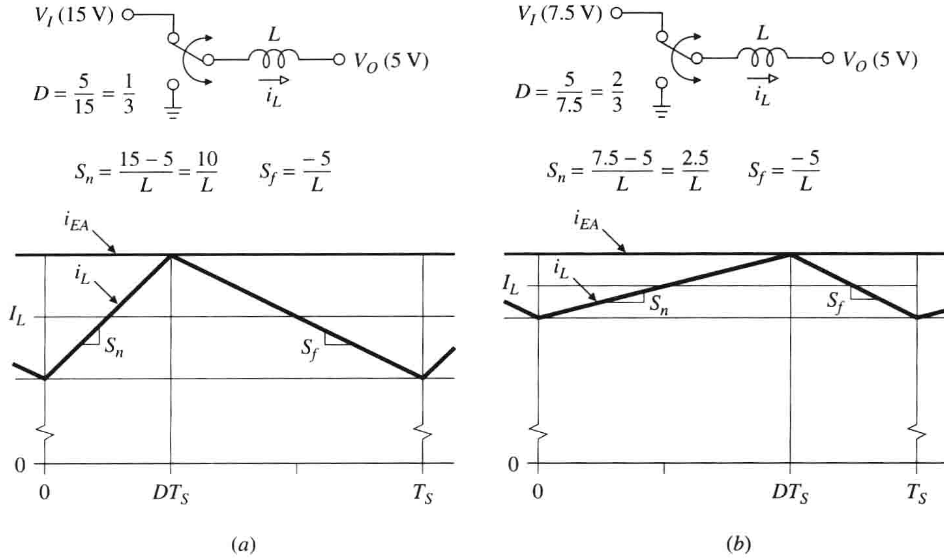


FIGURE 11.46
Without slope compensation, the average current I_L is strongly dependent on V_I.

equivalently, when i_L reaches the value v_{EA}/R_i. Denoting this value as

$$i_{EA} = \frac{v_{EA}}{R_i} \tag{11.55}$$

allows us to visualize a cycle exclusively in terms of currents as in Fig. 11.46 (for simplicity, the cycle is assumed to start at $t = 0$). This figure reveals the first of two possible flaws of uncompensated PCMC: the fact that if we lower/raise V_I while v_{EA} hasn't had time to change appreciably, the average inductor current I_L will raise/drop accordingly. This is so because the downslope $S_f(= -V_O/L)$ remains constant, so the upslope S_n must change to accommodate the change in V_I, and this in turn shifts I_L up or down. Since $I_O = I_L$, we have a case of very poor line regulation!

The second flaw is a form of instability known as *subharmonic oscillation,* which arises when $D > 0.5$ and the regulator is operating in the continuous mode, whereby i_L does not have time to drop to zero within a normal operating cycle. Refer to Fig. 11.47, which shows how an inductor current perturbation $i_l(0)$ at the beginning of a cycle (such a perturbation might be due, for instance, to a misfiring of the comparator in the course of the previous cycle) evolves into the perturbation $i_l(T_S)$ at the end of the cycle. Using simple geometry, we can write $i_l(0)/\Delta t = S_n$ and $i_l(T_S)/\Delta t = S_f$. Eliminating Δt gives

$$\frac{i_l(T_s)}{i_l(0)} = \frac{S_f}{S_n} = \frac{-D}{1-D} \tag{11.56}$$

indicating that (*a*) the polarity of $i_l(T_S)$ is *opposite* to that of $i_l(0)$, and (*b*) for $D < 0.5$ its magnitude will decrease (and die out after a sufficient number of cycles), but for $D > 0.5$ it will tend to increase from one cycle to the next, leading to the aforementioned subharmonic instability.

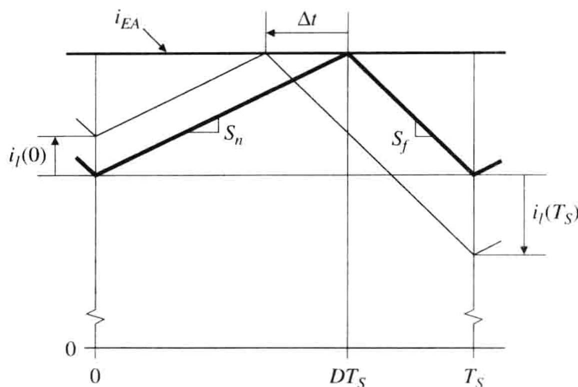


FIGURE 11.47
Illustrating subharmonic oscillation for $D > 0.5$.

Slope Compensation

Looking back at Fig. 11.46*b*, it seems that if we want it to retain the same I_L of Fig. 11.46*a*, we need to suitably "push down" its waveforms. We achieve this by suitably lowering $i_{L(pk)}$, a task in turn requiring that we suitably lower the value of i_{EA} responsible for tripping the comparator at DT_S. In fact, one can extend this line of reasoning over the entire range $0 < D < 1$, and use simple graphical techniques to figure that the locus of i_{EA} values needed to maintain I_L constant is itself a ramp! This locus, denoted as $i_{EA(\text{comp})}$ in Fig. 11.48, has a slope of $-S_f/2 = -V_O/(2L)$. The circuit of Fig. 11.44 synthesizes this locus by starting out with a voltage ramp v_{RAMP}, which it then subtracts from v_{EA} to produce the control voltage $v_{EA} - v_{RAMP}$.

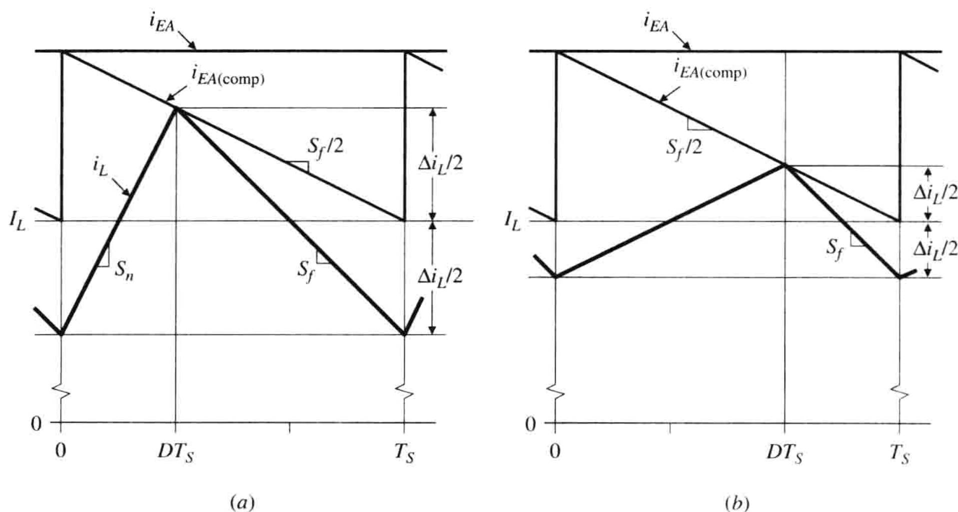


FIGURE 11.48
Showing how slope compensation maintains the same average current I_L regardless of variations in V_I and, hence, in D.

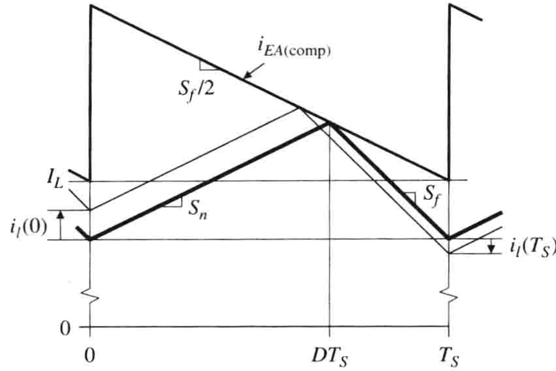


FIGURE 11.49
Slope compensation eliminates subharmonic oscillation regardless of D.

With this modification, aptly called *slope compensation*, Eq. (11.55) changes to

$$i_{EA(\text{comp})} = \frac{v_{EA} - v_{RAMP}}{R_i} \tag{11.57}$$

Comparison of Fig. 11.48 against Fig. 11.46 confirms the constancy of I_L in spite of the decrease in V_I and, hence, the increase in D as we go from (a) to (b).

But, there is more, because slope compensation also eliminates subharmonic oscillation, as depicted in Fig. 11.49. Using again graphical means, we observe that a beginning-of-cycle disturbance $i_l(0)$ will result in an end-of-cycle disturbance $i_l(T_S)$ of *lesser* magnitude, even though $D > 0.5$ (in fact, you can convince yourself that this holds also for the case $D < 0.5$ of Fig. 11.48a). Needless to say, slope compensation is a genial idea!

At this juncture it is worth pointing out that the slope compensation scheme of Fig. 11.44 is not unique. The frequently used alternative of Fig. 11.50 adds v_{RAMP} to $R_i i_L$ (instead of subtracting it from v_{EA}). The result, depicted in Fig. 11.51, is that now both slopes are augmented by the amount $S_e = -S_f/2 = V_O/(2L)$, $S_e > 0$, so the upslope is now steeper, and the downslope is flatter (from the viewpoint of the comparator, raising the i_L waveform as in Fig. 11.51 has the same effect as lowering the i_{EA} waveform as in Fig. 11.48).

The Control-to-Output Transfer Function

The ac analysis of a PCMC converter is rather complex because of the presence of two loops (an inner current loop to drive the inductor and an outer voltage loop to maintain voltage regulation). Moreover, since the inductor current is being sampled at each cycle, a rigorous analysis requires sampled-data theory.[13] Suffice it to say here that the *control-to-output* transfer function for the buck of Fig. 11.50 takes on the form[14,15]

$$H(j\omega) = \frac{V_o}{V_{ea}} = H_0 \frac{1 + j\omega/\omega_{\text{ESR}}}{(1 + j\omega/\omega_p)\left[1 - (\omega/0.5\omega_S)^2 + (j\omega/0.5\omega_S)/Q\right]} \tag{11.58}$$

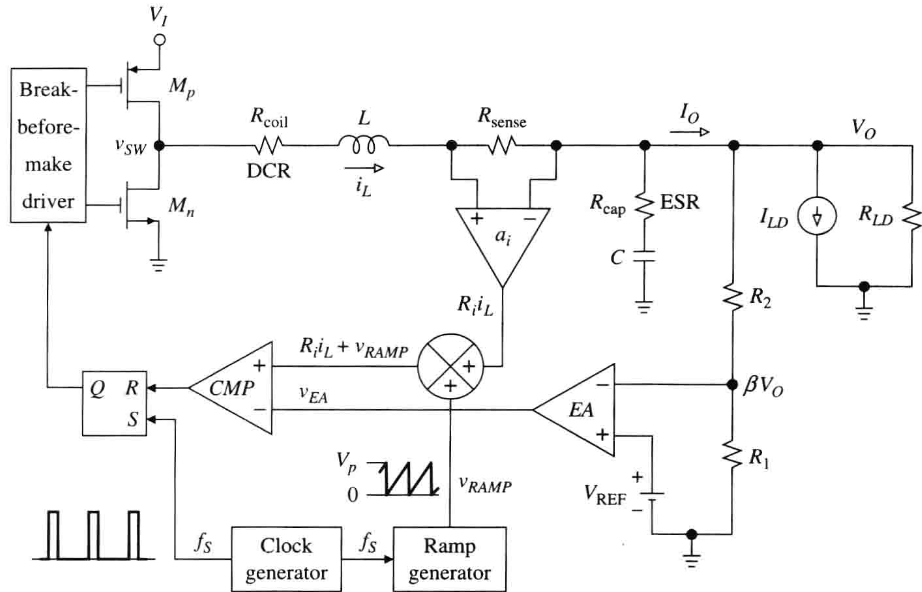


FIGURE 11.50
Alternative implementation of slope compensation, whereby v_{RAMP} is *added* to $R_i i_L$ (rather than subtracted from v_{EA}).

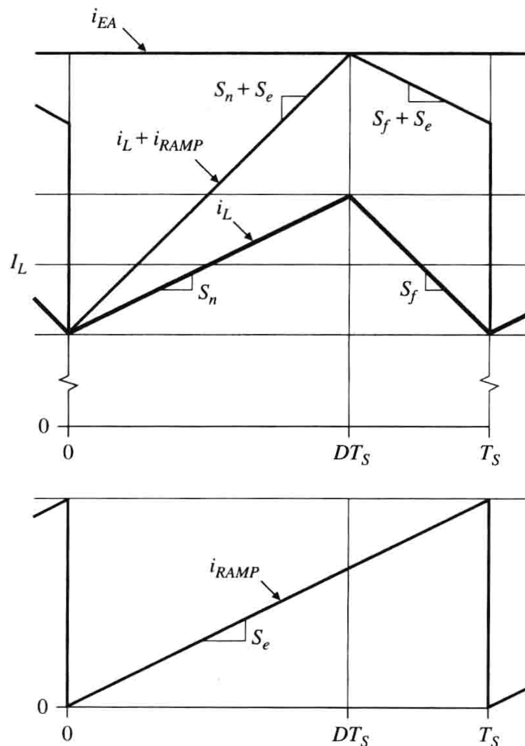


FIGURE 11.51
Waveforms for the alternative buck converter of Fig. 11.50.

where H_0 is the dc gain, ω_p is the dominant pole frequency established by the capacitor C, ω_{ESR} is the zero frequency due to the capacitor's ESR, and the second-order denominator term stems from the current sampling process at frequency ω_S. The following expressions[15,16] hold:

$$H_0 = \frac{K_M R_{LD}}{R_{LD} + R_{SW} + R_{\rm coil} + R_{\rm sense} + K_M R_i}$$

$$\omega_p = \frac{1}{[R_{LD} \| (K_M R_i)]C} \qquad \omega_{ESR} = \frac{1}{R_{\rm cap} C} \tag{11.59}$$

where R_{LD} is the resistive component of the load and K_M is a dimensionless factor

$$K_M = \frac{1}{\dfrac{(0.5 - D)R_i}{f_S L} + \dfrac{V_p}{V_I}} \tag{11.60}$$

and V_p is the peak amplitude of the compensation ramp. The Q-factor associated with the sampling process is[13]

$$Q = \frac{1/\pi}{(1 + S_e/S_n) \times (1 - D) - 0.5} \tag{11.61}$$

Note that in the absence of slope compensation ($S_n = 0$) and in the limit $D \to 0.5$ we get $Q \to \infty$. This explains the system's ability to sustain a subharmonic oscillation at half the sampling frequency, or $0.5 f_S$. For $D > 0.5$, Q becomes negative, indicating a growing oscillation. We now appreciate the need for slope compensation ($S_e > 0$) to avoid such instabilities. Beside the aforementioned choice of $S_e = -S_f/2$, another frequently used option is $S_e = -S_f = V_O/L$. With the latter option a disturbance i_l will get killed within a single cycle. Also referred to as *deadbeat* condition, this option gives, for $D = 0.5$, $Q = 2/\pi = 0.637$, which corresponds to a critically damped system.[14,15] Let us turn to an actual example to get a feel for the various parameters involved.

EXAMPLE 11.23. Let the buck converter of Fig. 11.50 have $V_I = 12$ V, $V_O = 5$ V, $I_O = 4$ A, $f_S = 500$ kHz, $V_p = 0.1$ V, $R_{\rm sense} = 20$ mΩ, $a_i = 10$ V/V, $L = 10\ \mu$H, $C = 40\ \mu$F, $R_{SW} = R_{\rm coil} = R_{\rm cap} = 10$ mΩ. (*a*) Calculate all relevant parameters for the case of a purely resistive load $R_{LD} = V_O/I_O$, which is typically the case when the regulator is tested in the lab. (*b*) Repeat, but for the opposite extreme of $R_{LD} \to \infty$, so $I_O = I_{LD}$ (a real-life load is likely to fall somewhere between these extremes). (*c*) Visualize the two frequency responses via PSpice, compare, and comment.

Solution. Using the above equations, calculate first the load-independent parameters:

$$R_i = 10 \times 20 \times 10^{-3} = 0.2\ \Omega \qquad D = 5/12 = 0.4167$$

$$S_n = (12 - 5)/(10 \times 10^{-6}) = 0.7\ \text{A}/\mu\text{s}$$

$$S_f = -5/(10 \times 10^{-6}) = -0.5\text{A}/\mu\text{s} \qquad S_e = -S_f/2 = 0.25\ \text{A}/\mu\text{s}$$

$$Q = \frac{1/\pi}{(1 + 0.25/0.7) \times (1 - 0.4167) - 0.5} = 1.09$$

$$K_M = \frac{1}{\dfrac{(0.5 - 0.4167)0.2}{500 \times 10^3 \times 10 \times 10^{-6}} + \dfrac{0.1}{12}} = 85.7 \qquad 0.5 f_S = 250\ \text{kHz}$$

$$f_{ESR} = \frac{1/(2\pi)}{0.01 \times 40 \times 10^{-6}} = 398\ \text{kHz}$$

(*a*) For the case of a purely resistive load, $R_{LD} = V_O/I_O = 1.25\ \Omega$, we have

$$H_0 = \frac{85.7 \times 1.25}{1.25 + 3 \times 0.01 + 0.02 + 85.7 \times 0.2} = 5.82$$

$$f_p = \frac{1/(2\pi)}{[1.25\|(85.7 \times 0.2)]40 \times 10^{-6}} = 3.42\ \text{kHz}$$

Consequently,

$$H(R_{LD} = 1.25\ \Omega) = 5.82 \frac{1 + \dfrac{jf}{398\ \text{kHz}}}{\left[1 + \dfrac{jf}{3.42\ \text{kHz}}\right]\left[1 - \left(\dfrac{f}{250\ \text{kHz}}\right)^2 + \dfrac{jf}{250\ \text{kHz}}\dfrac{1}{1.09}\right]}$$

(*b*) For the case of a purely current-sink load $I_{LD} = I_O$ in Fig. 11.50, we recalculate

$$\lim_{R_{LD}\to\infty} H_0 = K_M = 85.7 \qquad \lim_{R_{LD}\to\infty} f_p = \frac{1}{2\pi(K_M R_i)C} = 0.232\ \text{kHz}$$

so now

$$H(R_{LD} \to \infty) = 85.7 \frac{1 + \dfrac{jf}{398\ \text{kHz}}}{\left[1 + \dfrac{jf}{0.232\ \text{kHz}}\right]\left[1 - \left(\dfrac{f}{250\ \text{kHz}}\right)^2 + \dfrac{jf}{250\ \text{kHz}}\dfrac{1}{1.09}\right]}$$

(*c*) Using PSpice's Laplace blocks, we readily generate the plots of Fig. 11.52. The plots show the dominant pole, which is highest when R_{LD} is lowest, along with a slight amount of resonance ($Q \cong 1$) at half the clock frequency. Moreover, because of the zero frequency at f_{ESR}, the hig-frequency phase asymptote is $(-1-2+1) \times 90 = -180°$.

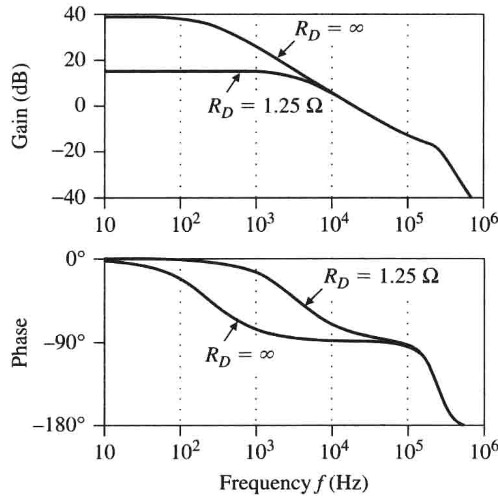


FIGURE 11.52
Bode plots of the control-to-output transfer function of the buck converter of Example 11.23 for the limiting cases $R_D = V_O/I_O = 1.25\ \Omega$ and $R_D \to \infty$.

Simplified AC Equivalent

We now wish to develop an intuitive feel for the control-to-output transfer function. So long as we use the proper amount of slope compensation and a low-ESR capacitor, for f well below $0.5 f_S$ and f_{ESR} we can approximate Eq. (11.58) as

$$H(jf) = \frac{V_o}{V_{ea}} \cong H_0 \frac{1}{(1 + jf/f_p)} \tag{11.62}$$

A well-designed converter will have $R_{SW} + R_{\text{coil}} + R_{\text{sense}} \ll R_{LD} + K_M R_i$, so Eq. (11.59) yields the simpler and more insightful expressions

$$H_0 \cong \frac{K_M R_{LD}}{R_{LD} + K_M R_i} = \frac{R_{LD} \| (K_M R_i)}{R_i} \qquad f_p = \frac{1}{2\pi [R_{LD} \| (K_M R_i)] C} \tag{11.63}$$

We now draw two important conclusions: (*a*) above f_p, yet well below both $0.5 f_S$ and f_{ESR}, the control-to-output transfer function exhibits a constant gain-bandwidth product of

$$\text{GBP} = H_0 \times f_p = \frac{1}{2\pi R_i C} \tag{11.64}$$

and (*b*) for $f \ll 0.5 f_S$ the inductor behaves like a VCCS with transconductance gain $1/R_i$ and parallel resistance $K_M R_i$. This results in the dramatic simplification of Fig. 11.53, which we expect to make the design of the compensator easier.

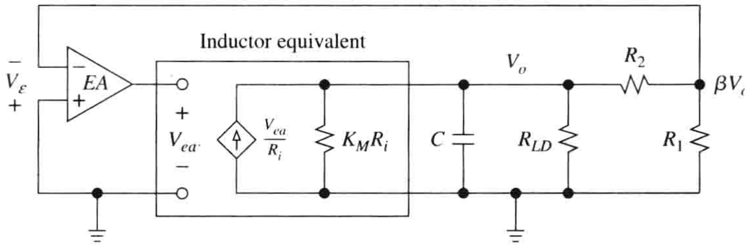


FIGURE 11.53
Simplified ac equivalent of the PCMC buck converter for $f \ll 0.5 f_S$ and $f \ll f_{\text{ESR}}$.

Error Amplifier Design

With the inductor delay removed from the loop, a Type 2 compensator will suffice. Monolithic converters usually implement the EA with an operational transconductance amplifier (OTA) terminated on an impedance designed to provide a Type 2 frequency profile. In Fig. 11.54 the OTA is modeled via a VCCS with transconductance gain G_m (typically on the order of 1 mA/V), output resistance R_o (typically on the order of 1 MΩ), and stray output-node capacitance C_o (typically in the range of a few pico-farads). The external components R_c and C_c establish the LHP zero frequency needed to boost the loop phase in the vicinity of the crossover frequency f_x, while C_{HFP} combines with C_o to establish a high-frequency pole needed to filter out high-frequency noise.

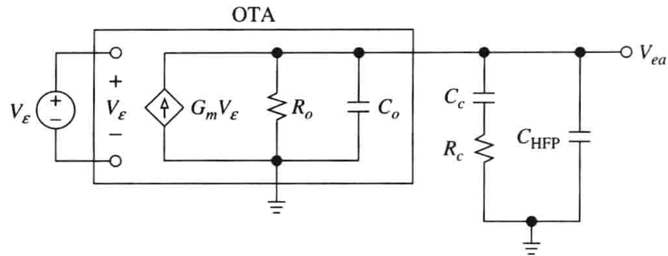


FIGURE 11.54
Using an OTA to implement a Type 2 error amplifier. The circuit is usually designed with $R_c \ll R_o$ and $C_c \gg (C_o + C_{\rm HFP})$.

Using physical inspection as in connection with Fig. 11.37, we readily construct the linearized Bode plot of Fig. 11.55, where

$$\omega_{p1} \cong \frac{1}{R_o C_c} \qquad \omega_{z1} \cong \frac{1}{R_c C_c} \qquad \omega_{p2} \cong \frac{1}{R_c(C_o + C_{\rm HFP})} \tag{11.65}$$

Note that R_c serves the dual purpose of establishing the zero frequency f_{z1} and the intermediate-frequency gain $G_m R_c$. Once G_m is known, the values of R_c, C_c, and $C_{\rm HFP}$ are uniquely determined by our choice of the crossover frequency as well as $\omega_{\rm HFP}$. A common strategy[16] is to make the first zero-pole pair of the compensator coincide with the first pole-zero pair of the control-to-output transfer function so as to attain a loop-gain rolloff of −20 dB/dec for a phase margin of about 90°. But, if necessary, the compensator can be suitably tweaked when the converter is tested in the lab.

EXAMPLE 11.24. Design an EA to stabilize the converter of Example 11.23 for the case $R_{LD} = V_O/I_O = 1.25\ \Omega$. Assume $V_{\rm REF} = 1.205$ V, $G_m = 0.8$ mA/V, $R_o = 1$ MΩ, and $C_o = 10$ pF. Verify with PSpice.

Solution. With reference to Fig. 11.44 we have $\beta = V_{\rm REF}/V_O = 1.205/5 = 1/4.15$, so $1 + R_2/R_1 = 4.15$. Pick $R_1 = 10$ kΩ and $R_2 = 31.5$ kΩ. Let us place the crossover

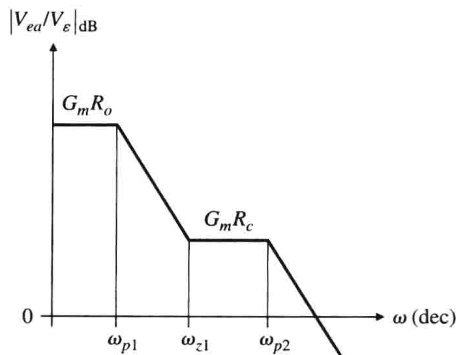


FIGURE 11.55
Linearized gain plot of the EA of Fig. 11.54.

frequency at $f_x = f_S/10 = 500/10 = 50$ kHz, where we have

$$|H(jf_x)| = \frac{\text{GBP}}{f_x} = \frac{1/(2\pi R_i C)}{f_x} = \frac{1/(2\pi \times 0.2 \times 40 \times 10^{-6})}{50 \times 10^3} = \frac{1}{2.51}$$

To produce $|T(jf_x)| = 1$, the EA's gain at f_x must make up for the combined attenuations due to β and $|H(jf_x)|$. Since f_x is going to lie between f_{z1} and f_{p2}, where the EA gain is flat, we impose

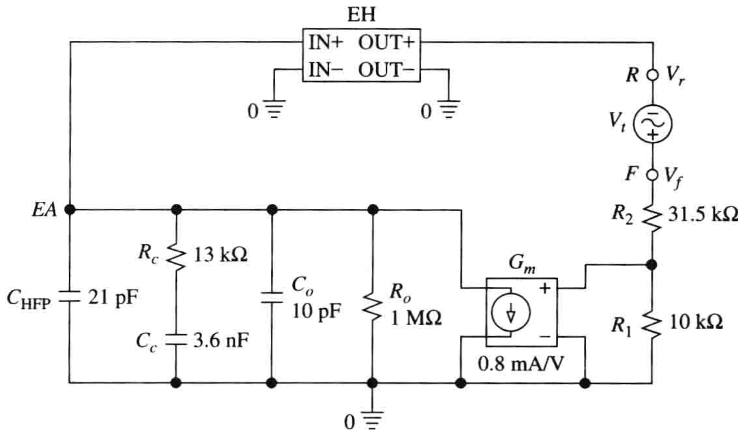
$$G_m R_c = 4.15 \times 2.51 = 10.4 \text{ V/V}$$

and get $R_c = 10.4/(0.8 \times 10^{-3}) = 13$ kΩ. Next, impose $f_{z1} = f_p$ for the first zero-pole cancellation,

$$\frac{1}{2\pi \times 13 \times 10^3 \times C_c} = 3.42 \times 10^3$$

and get $C_c \cong 3.6$ nF. Finally, impose $f_{p2} = f_{\text{ESR}}$ for the second pole-zero cancellation,

$$\frac{1}{2\pi \times 13 \times 10^3 \times (C_o + C_{\text{HFP}})} = 398 \times 10^3$$



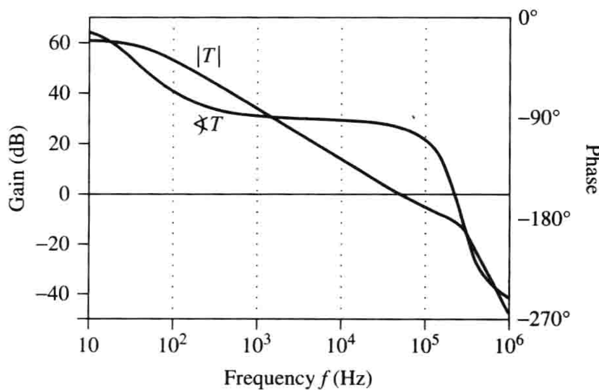


FIGURE 11.56
PSpice simulation of the buck converter of Example 11.24.

Substituting $C_o = 10$ pF gives $C_{\text{HFP}} \cong 21$ pF. Finally, we use the voltage injection technique depicted in Fig. 11.56 to plot the loop gain (note the constant rolloff of -20 dB/dec, as expected). The crossover frequency is $f_x = 50.1$ kHz, where $\angle T = -98.4°$. Consequently, $\phi_m = 180 - 98.4 = 81.6°$.

11.10 PCMC OF BOOST CONVERTERS

Figure 11.57 shows the peak current mode control (PCMC) for the boost topology. As depicted in the timing diagram of Fig. 11.58, a cycle initiates when a clock pulse sets the flip-flop. This turns on the switch M, causing i_L to ramp up with a slope of $S_n = V_I/L$. Note that during this part of the cycle the diode is cut off, so the load is feeding entirely off the capacitor.

Once i_L reaches a peak value established by the error amplifier EA according to the slope compensation mechanism, the comparator fires, resetting the flip-flop and turning off the switch M. At this point the magnetic field inside the inductor starts to collapse, so v_L changes polarity and increases in magnitude until it turns on the Schottky diode. Assuming the diode drops zero voltage for simplicity, we can say that the voltage across the inductance is now $v_L = V_I - V_O$, $v_L < 0$, so i_L ramps down with a slope of $S_f = (V_I - V_O)/L$.

Initially, following power turn-on, the value of i_L at the end of each cycle exceeds that at the beginning of the cycle in order to build up enough average current in the inductor to satisfy the current demanded by the load. After a sufficient number of

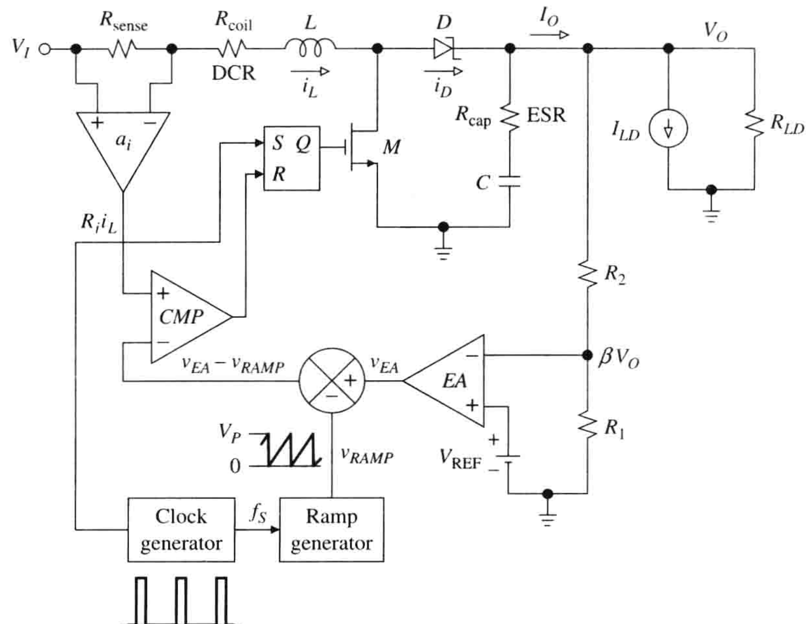


FIGURE 11.57
Boost converter with peak current mode control (PCMC).

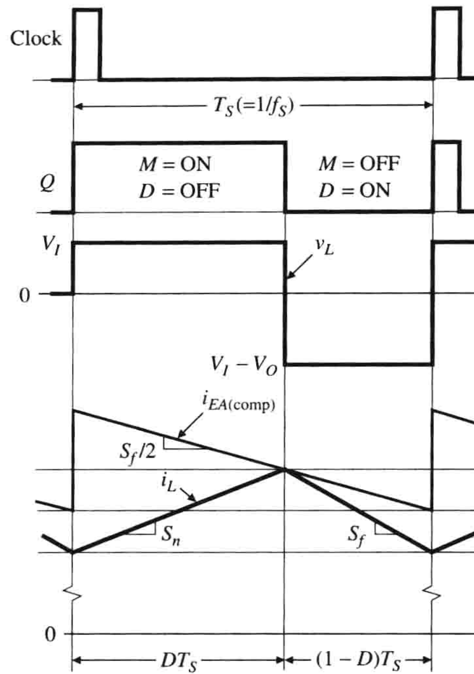


FIGURE 11.58
Timing of PCM control for the buck converter.

cycles, a steady-state condition is reached, whereby the end-of-cycle value of i_L will equal the beginning-of-cycle value. Alternatively, we invoke the inductance volt-second balance condition by saying that in steady state v_L must average to zero, though it continues to alternate between V_I and $V_I - V_O$, as shown. As discussed in Section 11.6, a boost converter gives

$$V_O = \frac{1}{1-D} V_I \qquad I_L = \frac{1}{1-D} I_O \tag{11.66}$$

The Right-Half-Plane Zero (RHPZ)

As mentioned, during the first part of the cycle, when the diode is off, the load is feeding entirely off the capacitor (this is in marked contrast with the buck converter, where the capacitor/load combination receives current via the inductor during the entire cycle). We shall now demonstrate that this feature tends to pose stability issues. Consider first the equilibrium situation of Fig. 11.59a. By the capacitance ampere-second balance condition, the capacitance current in Fig. 11.57 must average zero, so by KCL the average current I_O delivered to the load must equal the average current I_D supplied via the diode, or

$$I_O = I_D \tag{11.67}$$

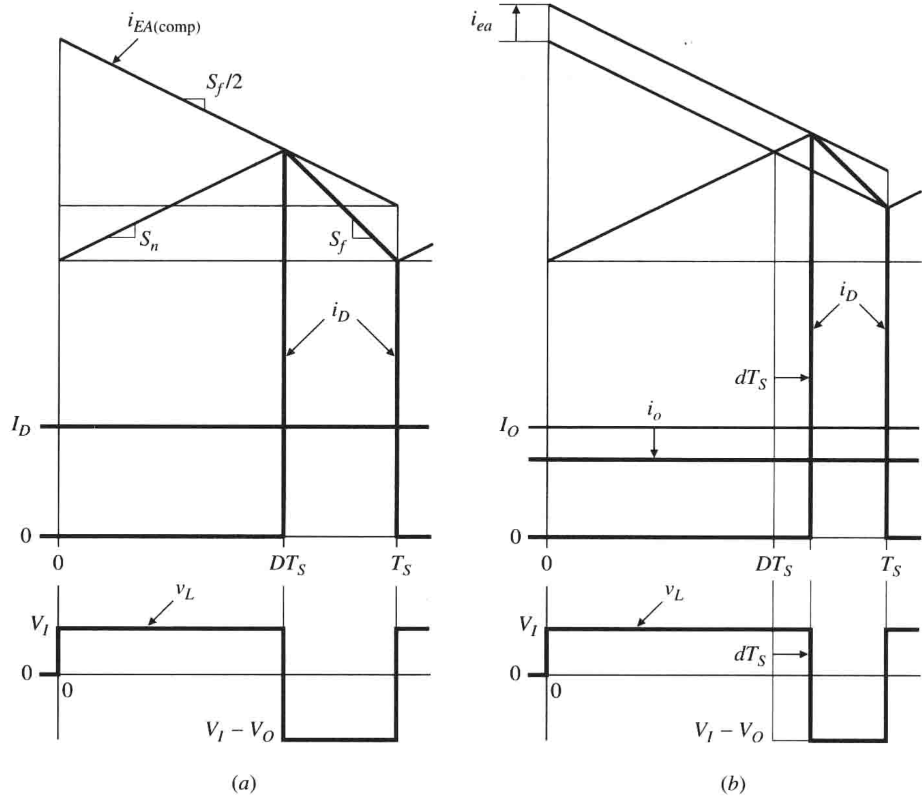


FIGURE 11.59
Investigating the physical origin of the right-half-plane zero (RHPZ).

(You can visualize I_D as the area under the i_D waveform, divided by T_S.) Now suppose that in response to a jump in the current demand by the load, the EA raises its output voltage v_{EA} by an amount v_{ea}. This, in turn, up-shifts the waveform $i_{EA(\text{comp})}$ by the amount

$$i_{ea} = \frac{v_{ea}}{R_i}$$

As depicted in Fig. 11.59*b*, this upshift delays the flip-flop resetting by dT_S. This delay, while desirable because it allows for more energy to build up in the inductor in response to the jump in the load demand, is also unwelcome because it reduces the conduction time of the diode. Since I_D is the area under the i_D waveform divided by T_S, a reduced conduction time implies a drop in I_D and, hence, a drop i_o in the average current $I_O(= I_D)$ passed on to the load. In summary, instead of opposing the load disturbance, the circuit reinforces it (at least initially), as if negative feedback had turned positive! Mathematically, this phenomenon is at the basis of the much-dreaded right-half-plane zero (RHPZ) of boost-type converters operating in the continuous conduction mode.

For a mathematical analysis of the RHPZ, start out with the equilibrium situation of Fig. 11.59a, for which we have, by the definition of average,

$$V_L = \frac{DT_S V_I}{T_S} + \frac{(1-D)T_S(V_I - V_O)}{T_S} = V_I - V_O(1-D)$$

Though $V_L = 0$ in equilibrium, V_L will increase from 0 to $0 + v_l$ if the duty cycle is increased from D to $D + d$ as in Fig. 11.59b. So long as V_I and V_O don't change appreciably from one cycle to the next, the above expression gives $v_l = 0 - V_O(0 - d)$, or

$$v_l = V_O d$$

Similarly, the condition $I_O = (1-D)I_L$ of Eq. (11.66) holds in equilibrium, but as we raise the duty cycle to $D + d$, it changes as

$$I_O + i_o = [1 - (D+d)] \times (I_L + i_l) = (1-D)I_L + (1-D)i_l - I_L d + i_l d$$
$$\cong I_O + (1-D)i_l - I_L d$$

where the product of small terms $i_l \times d$ has been ignored. Clearly we have

$$i_o \cong (1-D)i_l - I_L d$$

Substituting $i_l = v_l/(sL) = V_O d/(sL)$ and collecting gives

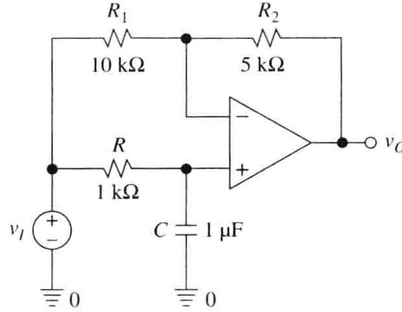
$$i_o = \left[\frac{(1-D)V_O}{sL} - I_L\right] d$$

It is apparent that i_o vanishes for $s = (1-D)V_O/(LI_L)$, $s > 0$, indicating a zero in the right half of the complex plane. Letting $I_L = I_O/(1-D)$, we express the RHPZ in the more common form

$$\omega_{\text{RHPZ}} = \frac{(1-D)^2 V_O}{I_O L} \tag{11.68}$$

As we shall see shortly, this root results in a numerator term of the type $(1 - j\omega/\omega_{\text{RHPZ}})$ in the control-to-output transfer function V_o/V_{ea}. Subjecting the system to a step change in v_{ea} will elicit, at least initially, a response i_o in the *opposite* direction of the intended value. This, because the term $-j\omega/\omega_{\text{RHPZ}}$ prevails at high frequencies, which is where the energy of the step's leading edge is concentrated. However, if we allow v_{ea} to change at a suitably *slow rate* so as to render $-j\omega/\omega_{\text{RHPZ}}$ negligible compared to unity, the system will respond in the intended direction almost right from the beginning.

Let us illustrate the above concepts via a more familiar circuit such as that of Fig. 11.60a. For $R_2 = 0$ the circuit is a single-pole RC network with the op amp acting as a buffer (R_1 plays no role in this case), so the step response is the familiar exponential transient depicted as a thin trace in Fig. 11.60b (top). However, with $R_2 \neq 0$, the circuit exhibits also an RHPZ (see Problem 11.48), so its step response starts out with a 0.5-V swing in the *opposite* direction, though eventually it will settle to +1V as intended (can you justify this physically in terms of capacitor and op amp behavior?). If, instead of an abrupt step, we apply a gradual ramp as in Fig. 11.60b (bottom), the capacitor will be given more time to charge up, resulting in a less pronounced opposing swing. In terms of the converter's control-to-output transfer



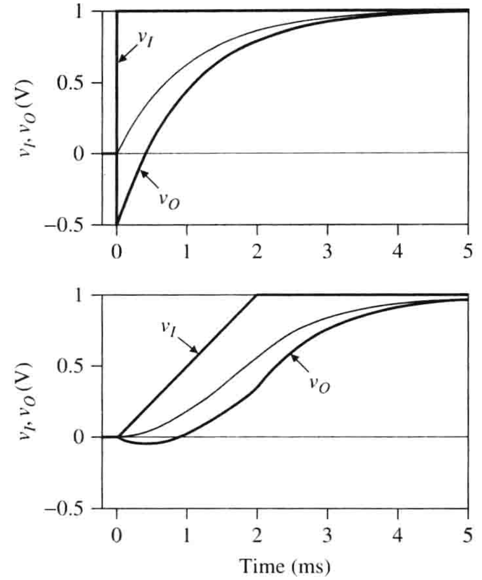


FIGURE 11.60
(*a*) Circuit with a LHP pole and a RHP zero, and (*b*) its step response (top) and ramp response (bottom). The thin traces refer to the case $R_2 = 0$.

function, this means that we can avoid the destabilizing tendency of the RHPZ by making sure that the loop gain T at ω_{RHPZ} has dropped to a suitably small value. Equivalently, this requires placing the crossover frequency *well below* the RHP zero frequency, such as $\omega_x \leq \omega_{\text{RHPZ}}/10$.

The Control-to-Output Transfer Function

The control-to-output transfer function of the boost converter takes on the form[15,16]

$$H(j\omega) = \frac{V_o}{V_{ea}} = H_0 \frac{(1 - j\omega/\omega_{\text{RHPZ}})(1 + j\omega/\omega_{\text{ESR}})}{(1 + j\omega/\omega_p)[1 - (\omega/0.5\omega_S)^2 + (j\omega/0.5\omega_S)/Q]} \tag{11.69}$$

where H_0 is the dc gain, ω_p is the dominant-pole frequency established by the output capacitor C, ω_{RHPZ} is the aforementioned RHPZ frequency exhibited by boost converters operating in the continuos-current mode, ω_{ESR} is the zero frequency due to the capacitor's ESR, and the second-order denominator term stems from the current sampling process at frequency ω_S. The following expressions hold:[15]

$$\omega_{\text{RHPZ}} = \frac{(1-D)^2 V_O}{I_O L} \qquad \omega_{\text{ESR}} = \frac{1}{R_{\text{cap}} C} \qquad Q = \frac{1/\pi}{(1 + S_e/S_n) \times (1 - D) - 0.5} \tag{11.70}$$

The dc gain H_0 and the dominant pole f_p are such that *above* f_p but *well below* f_{RHPZ}, f_{ESR}, and $0.5 f_S$ the gain-bandwidth product is constant and such that

$$\mathrm{GBP} = H_0 \times f_p = \frac{1 - D}{2\pi R_i C} \tag{11.71}$$

Except for the numerator term $(1 - D)$, this is similar to the buck case, so we can again exploit this feature to facilitate the design of the EA.

EXAMPLE 11.25. Using an OTA with $G_m = 1$ mA/V, $R_o = 1$ MΩ, and $C_o = 5$ pF, design an EA to stabilize a boost converter having $V_I = 5$ V, $V_O = 12$ V, $I_O = 4$ A, $V_{REF} = 1.205$ V, $f_S = 400$ kHz, $R_i = 0.1$ Ω, $L = 10$ μH, $R_{coil} = 10$ mΩ, $C = 200$ μF, and $R_{cap} = 5$ mΩ. Assume the converter is terminated on a purely resistive load $R_{LD} = V_O/I_O = 3$ Ω, in which case the dc gain takes on the form[15]

$$H_0 \cong \frac{R_{LD}(1 - D)}{2R_i} \tag{11.72}$$

What is the phase margin of the converter?

Solution. With reference to Fig. 11.57 we have $\beta = V_{REF}/V_O = 1.205/12 = 1/9.96$, so $R_2/R_1 = 8.96$. Pick $R_1 = 10$ kΩ and $R_2 = 89.6$ kΩ. For this regulator we have $D = 1 - V_I/V_O = 1 - 5/12 = 0.583$, so using the above equations we find

$$H_0 = 6.25 \quad f_p = 0.53 \text{ kHz} \quad f_{RHPZ} = 8.3 \text{ kHz} \quad f_{ESR} = 159 \text{ kHz} \qquad Q = 1.53$$

One possible strategy[15] is to set the crossover frequency f_x equal to the lower of either $(1/4) f_{RHPZ}$ or $(1/10) f_S$. So, in the present case let $f_x = 8.3/4 \cong 2.0$ kHz. Next, calculate the gain required of the EA at f_x. Since f_x is much lower than both $0.5 f_S$ and f_{ESR}, we drop the corresponding terms in Eq. (11.69) and write

$$|H(jf_x)| \cong H_0 \sqrt{\frac{1 + (f_x/f_{RHPZ})^2}{1 + (f_x/f_p)^2}} = 6.25 \sqrt{\frac{1 + (2.0/8.3)^2}{1 + (2.0/0.53)^2}} \cong 1.7$$

To produce $|T(jf_x)| = 1$, the EA gain at f_x must compensate for the combined attenuation due to β and the gain of 1.7 due to H. Since f_x is going to lie between f_{z1} and f_{p2}, where the EA gain is flat, we impose

$$G_m R_c = 9.96/1.7 = 5.8 \text{ V/V}$$

This gives $R_c = 5.8/(1 \times 10^{-3}) = 5.8$ kΩ. Next, impose[15] $f_{z1} = f_x/2.5 = 2.0/2.5 = 0.8$ kHz, which should still give a reasonable phase margin (if subsequent lab measurements prove this choice to be inadequate, we can always lower or raise f_{z1} by raising or lowering C_c). Thus, letting

$$\frac{1}{2\pi (5.8 \text{ k}\Omega) C_c} = 0.8 \text{ kHz}$$

gives $C_c \cong 34$ nF. Finally, make the EA's high-frequency pole f_{p2} equal[15] to the lower of either f_{RHPZ} or f_{ESR} so as to ensure proper frequency rolloff for T. In the present case let $f_{p2} = f_{RHPZ}$, or

$$\frac{1}{2\pi (5.8 \text{ k}\Omega) \times (C_o + C_{HFP})} = 8.3 \text{ kHz}$$

Letting $C_o = 5$ pF gives $C_{HFP} = 3.3$ nF. The simulation of Fig. 11.61 gives $f_x = 1.87$ kHz and $\phi_m \cong 59°$.

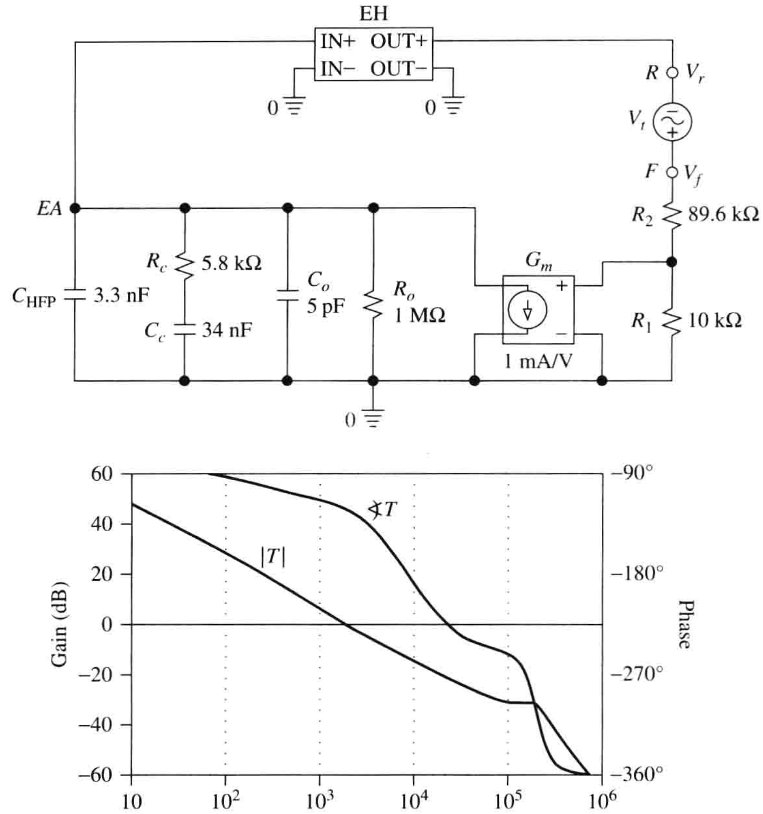


FIGURE 11.61
PSpice simulation of the boost converter of Example 11.25.

PROBLEMS

11.1 Performance specifications

11.1 An unregulated voltage $V_I = (26 \pm 2)$ V is applied to a shunt regulator consisting of a series 200-Ω resistor and an 18-V, 20-Ω shunt diode. The output of this regulator is then fed to a second regulator consisting of a 300-Ω series resistor and a 12-V, 10-Ω shunt diode to achieve an even better regulated voltage V_O for a load R_L. Sketch the circuit; then, find its line and load regulation and the minimum R_L allowed.

11.2 Using a 6.2-V Zener diode and a 741 op amp, design a negative self-biased reference that accepts an unregulated negative voltage V_I and gives a regulated output V_O adjustable from -10 V to -15 V by means of a 10-kΩ pot. What are the permissible ranges for V_I and I_O?

11.3 At $I_Z = 7.5$ mA the 1N827 thermally compensated Zener diode gives $V_Z = 6.2$ V ± 5% and TC(V_Z) = 10 ppm/°C. (*a*) Using this diode, along with an op amp having TC(V_{OS}) = 5 μV/°C, design a 10.0-V self-regulated reference with provision for the exact adjustment of V_O. (*b*) Estimate the worst-case change in V_O for a temperature variation of 0 °C to 70 °C.

11.4 Consider the circuit obtained from the self-regulated reference of Fig. 11.4 by lifting the left terminals of R_1 and D_z off ground, connecting them together, and then returning the resulting common node to ground via a variable resistance R. (*a*) Show that this modification allows us to vary V_O without altering the diode current. (*b*) Obtain a relationship between V_O and V_Z. (*c*) Specify standard components for a variable reference from 10 V to 20 V using the 1N827 diode of Problem 11.3 as the reference element.

11.5 (*a*) Assuming matched BJTs in the alternative bandgap cell[2] of Fig. P11.5, show that $V_{REF} = V_{BE1} + KV_T$, $K = (R_2/R_3)\ln(R_2/R_1)$. (*b*) Assuming I_s (25 °C) = 5×10^{-15} A for both BJTs, specify suitable components for TC(V_{REF}) = 0 at 25 °C.

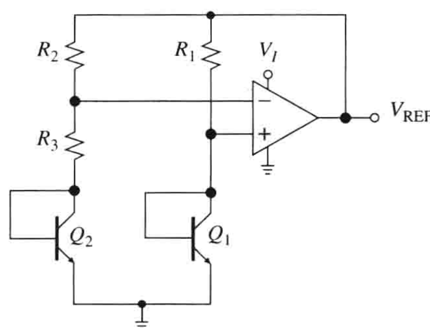


FIGURE P11.5

11.6 The bandgap reference[2] of Fig. P11.6 is known as the *Widlar bandgap cell* for its inventor. (*a*) Assuming matched BJTs with negligible base currents, show that $V_{REF} = V_{BE3} + KV_T$, $K = (R_2/R_3)\ln(I_{C1}/I_{C2})$. (*b*) Specify suitable components for TC(V_{REF}) = 0 at 25 °C if I_s (25 °C) = 2×10^{-15} A for all BJTs, $I_{C1} = I_{C3} = 0.2$ mA, and $I_{C2} = I_{C1}/5$.

11.7 Prove Eq. (11.17). Hence, assuming $m_p = 10$ and $V_T = 26$ mV, specify suitable resistance values for a current of 100 μA through Q_1.

11.3 Voltage-reference applications

11.8 Suppose you have (1) an op amp whose input voltage range and output voltage swing extend within 1 V of each supply; (2) a 5.0-V voltage reference of the type of Fig. 11.1 with a dropout voltage of 1 V; (3) a pair of 10-kΩ resistors. (*a*) Using only the above components, design a circuit to synthesize $V_{O1} = +5$ V and $V_{O2} = -5$ V. What are the minimum supply voltages required to power your circuit? (*b*) Repeat, but for $V_{O1} = +2.5$ V and $V_{O2} = +5$ V. (*c*) Repeat, but for $V_{O1} = +5$ V and $V_{O2} = +10$ V. (*d*) Repeat, but for $V_{O1} = +2.5$ V and $V_{O2} = -2.5$ V.

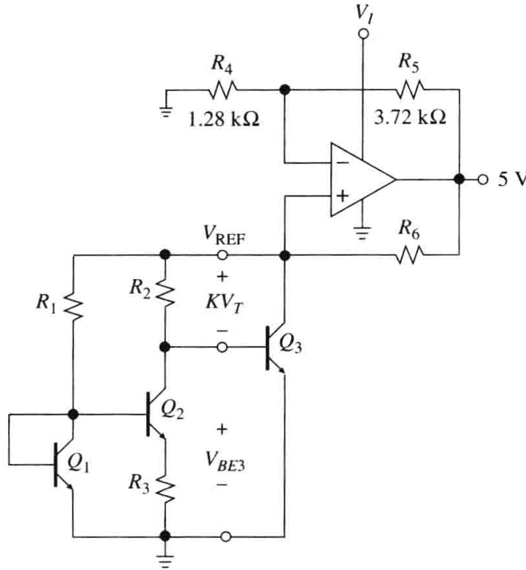


FIGURE P11.6

11.9 (*a*) Using a 5-V voltage reference, an op amp, a 10-kΩ potentiometer, and resistors as needed, design a variable voltage reference over the range $-5\text{ V} \le V_O \le +5\text{ V}$. (*b*) Repeat, but for $-10\text{ V} \le V_O \le +10\text{ V}$.

11.10 The LT1029 is a 5-V reference diode that operates with any current between 0.6 mA and 10 mA, and has a maximum TC of 20 ppm/°C. Using the LT1029 and a JFET-input op amp with $\text{TC}(V_{OS}) = 6\ \mu\text{V}/^\circ\text{C}$, design ±2.5-V split references and estimate their worst-case thermal drifts. Assume ±5-V power supplies.

11.11 Assuming ±15-V supplies and using an LM385 2.5-V diode with a bias current of 100 μA as a voltage reference, design a current generator whose output is variable over the range $-1\text{ mA} \le I_O \le 1\text{ mA}$ by means of a 10-kΩ pot.

11.12 The LM10 consists of two op amps and a 200-mV reference internally connected as in Fig. P11.12. The op amps have rail-to-rail output swing capability, and the device draws a maximum quiescent current of 0.5 mA from a supply voltage anywhere between 1.1 V and 40 V. The LM10C version has $\text{TC}(V_{\text{REF}}) = 0.003\ \%/^\circ\text{C}$, $\text{TC}(V_{OS}) = 5\ \mu\text{V}/^\circ\text{C}$, line regulation = 0.0001 %/V, and $\text{CMRR}_{\text{dB}} \cong \text{PSRR}_{\text{dB}} \cong 90$ dB. (*a*) Using the

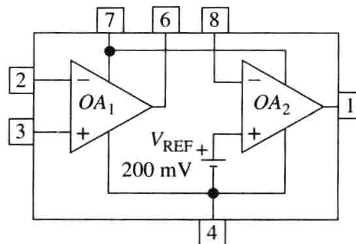


FIGURE P11.12

LM10C, design a voltage reference continuously variable from 0 to 10 V by means of a 10-kΩ pot. (*b*) Find the worst-case thermal drift and line regulation of your circuit.

11.13 Figure P11.13 shows a current generator using the LM10 of Problem 11.12. (*a*) Analyze the circuit and show that as long as the externally applied voltage between its terminals is sufficient to keep the op amps in the linear region, the current that the circuit sinks at the positive terminal and sources at the negative terminal is $I_O = (1 + R_2/R_3)V_{REF}/R_1$. (*b*) Specify suitable components for $I_O = 5$ mA. (*c*) What is the range of external voltages over which your circuit will operate properly?

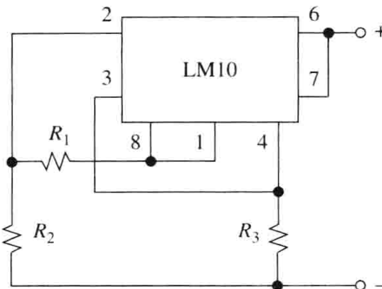


FIGURE P11.13

11.14 A 5-V voltage reference having output noise density $e_{nw1} = 100$ nV/$\sqrt{\text{Hz}}$ is buffered by a voltage follower implemented with a 1-MHz JFET-input op amp having input noise density of $e_{nw2} = 25$ nV/$\sqrt{\text{Hz}}$. (*a*) What is the total output noise E_{no}? (*b*) Modify your follower in the manner of Fig. 7.13 so as to lower E_{no} to 1/10 (or less) of its original value.

11.15 Design a circuit that senses the temperatures T_1 and T_2 at two different sites, and yields $V_O = (0.1\text{ V})(T_2 - T_1)$, T_1 and T_2 in degrees Celsius. The circuit uses two matched diodes with I_s (25 °C) = 2 fA as temperature sensors, and two potentiometers for its calibration. Describe the calibration procedure.

11.16 Specify suitable components in the circuit of Fig. 11.14 for a Fahrenheit sensor with a sensitivity of 10 mV/°F. Outline its calibration.

11.4 Linear regulators

11.17 Using a 741 op amp, an LM385 2.5-V reference diode, and *pnp* BJTs, design an overload-protected negative regulator with $V_O = -12$ V and $I_{O(\text{max})} = 100$ mA.

11.18 Using the LM10 of Problem 11.12 and two *npn* BJTs, design a 100-mA overload-protected voltage regulator whose output can be varied from 0 to 15 V by means of a 10-kΩ pot. Show how you power your circuit, and estimate the lowest permissible supply voltage.

11.19 Figure P11.19 shows a high-voltage regulator based on the LM10 of Problem 11.12. Since the LM10 is powered by three V_{BE} drops, the high-voltage capabilities of the circuit are limited only by the external components. (*a*) Analyze the circuit and find V_O in terms of V_{REF}. (*b*) Specify R_1 and R_2 for $V_O = 100$ V. (*c*) Assuming typical BJT parametrs, estimate V_{DO} for $I_O = 1$ A.

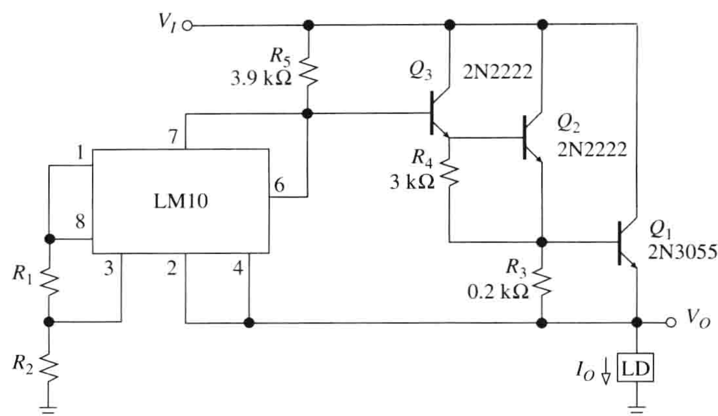


FIGURE P11.19

11.5 Linear-regulator applications

11.20 The LM338 is a 1.2-V, 5-A, adjustable regulator having $V_{\mathrm{DO}} = 2.5$ V, a maximum input-output differential voltage of 35 V, and an adjustment-pin current of 45 μA. Using the LM338, design a 5-A regulator whose output can be varied from 0 V to 5 V via a 10-kΩ pot. What are the power-supply requirements of your circuit?

11.21 Using the LT1029 reference diode of Problem 11.10 and the LM338 voltage regulator of Problem 11.20, design a minimum-component circuit for the simultaneous generation of a 5-V reference voltage and a 15-V, 5-A supply voltage. What is the permissible unregulated input voltage range of your circuit?

11.22 Using a μA7805 5-V regulator and 0.25-W (or less) resistances, design a 1-A current source. What is its voltage compliance as a function of the supply voltage?

11.23 In Fig. 11.23*a* let the common terminal of the regulator be connected directly to the node shared by R_1 and R_2 to save the op amp. Assuming a μA7805 5-V regulator, whose quiescent current is rated as $I_Q \cong 4$ mA flowing out of the common terminal, find suitable resistances for $V_O = 12$ V; then find the permissible range for V_{CC}, as well as the load and line regulation.

11.24 In the circuit of Fig. 11.19*b* let $V_I = 25$ V and $R_1 = 2.5$ Ω, and let R_2 be an arbitrary load. Find the Norton equivalent of the circuit seen by the load, along with its voltage compliance, given the following LM317 specifications: $V_{\mathrm{DO}} \cong 2$ V, line regulation = 0.07%/V maximum, and $\Delta I_{\mathrm{ADJ}} = 5$ μA maximum for $2.5 \text{ V} \le (V_I - V_O) \le 40$ V.

11.25 The LT337 is a −1.25-V, 1.5-A, adjustable negative regulator with $\Delta V_{\mathrm{REG}}/\Delta(V_I - V_O) = 0.03\%/\text{V}$ maximum, and $\Delta I_{\mathrm{ADJ}}/\Delta(V_I - V_O) = 0.135$ μA/V maximum. Using this device, design a 500-mA current sink; next, find its Norton equivalent.

11.26 Using an LM317 1.25-V, adjustable positive regulator, and an LM337 −1.25-V, adjustable negative regulator, design a dual-tracking bench power supply whose outputs are adjustable from ±1.25 V to ±20 V by means of a single 10-kΩ pot. See Problems 11.24 and 11.25 for the specifications of these regulators.

11.27 (*a*) Find the maximum allowable operating ambient temperature if $T_{J(\max)} = 190\ ^\circ\text{C}$, $P_{D(\max)} = 1$ W, and $\theta_{JA} = 60\ ^\circ$C/W. (*b*) Find θ_{JA} for a 5-V regulator with $T_{J(\max)} = 150\ ^\circ$C to deliver 1 A at $V_I = 10$ V and $T_A = 50\ ^\circ$C. Can a μA7805 operating in free air do it?

11.28 In the circuit of Fig. 11.23*b* the pot is replaced by the series combination of a 2-kΩ resistance between the inverting input and the regulator's output, and an 18-kΩ resistance between the inverting input and the regulator's common. Assuming $\pm$18-V supplies, $R = 1.00\ \Omega$, and a μA7805 regulator in the TO-220 package, specify a heat sink for operation all the way down to a load voltage of 0 V with $T_{A(\max)} = 60\ ^\circ$C.

11.29 Using the LM10 of Problem 11.12 and a 1.5-V, 2-mA LED, design an indicator circuit that monitors its own power supply and turns the LED off whenever the supply drops below 4.75 V.

11.30 Specify components in the circuit of Fig. 11.29*a* to provide OV protection when V_{CC} tries to rise above 6.5 V, and issue a $\overline{\text{PFAIL}}$ command when the 120-V (rms), 60-Hz ac line tries to drop below 80% of its nominal value.

11.6 Switching regulators

11.31 The switched coil of Fig. 11.31*e* bears some similarity to the switched capacitor of Fig. 4.23*a*. (*a*) Assuming $V_{\text{SAT}} = V_F = 0$, compare the two arrangements and point out similarities as well as fundamental differences. (*b*) Assuming the coil current waveform of Fig. 11.33*a*, show that the power transferred by the coil from V_I to V_O is $P = f_S W_{\text{cycle}}$, where $W_{\text{cycle}} = L I_L \Delta i_L$ is the energy packet transferred during each cycle.

11.32 (*a*) Derive Eq. (11.40). Then, assuming $I_O = 1$ A and $\Delta i_L = 0.2$ A, estimate I_p as well as the minimum value of I_O for continuous operation for the case of (*b*) a buck converter with $V_I = 12$ V and $V_O = 5$ V, (*c*) a boost converter with $V_I = 5$ V and $V_O = 12$ V, and (*d*) an inverting converter with $V_I = 5$ V and $V_O = -15$ V.

11.33 An inverting converter with $5\text{ V} \le V_I \le 10$ V is to deliver $V_O = -12$ V at a full load of 1 A. Assuming continuous operation with $V_{\text{SAT}} = V_F = 0.5$ V, find the required range for D, as well as the maximum value of I_I.

11.34 A buck-boost converter is powered from +15 V and operates at 150 kHz. Specify L, C, and ESR for $V_O = -15$ V, $V_{ro(\max)} = 150$ mV, and continuous-mode operation over the range $0.2\text{ A} \le I_O \le 1$ A.

11.35 A buck converter has $V_I = 20$ V, $V_O = 5$ V, $f_S = 100$ kHz, $L = 50\ \mu$H, and $C = 500\ \mu$F. Assuming $V_{\text{SAT}} = V_F = 0$ and ESR $= 0$ sketch and label i_{SW}, i_D, i_L, i_C, and the voltage v_X at the left terminal of L for the case of (*a*) continuous-mode operation with $I_O = 3$ A, and (*b*) discontinuous-mode operation with $t_{\text{ON}} = 2\ \mu$s.

11.36 Discuss how η in the converter of Example 11.20 is affected by (*a*) doubling V_I, and (*b*) doubling f_S.

11.7 The error amplifier

11.37 (*a*) Assuming ideal op amp in the EA of Fig. 11.36, expand $Z_A(j\omega)$ and $Z_B(j\omega)$ to obtain Eqs. (11.47) and (11.49). (*b*) Specify suitable component values for $f_{z1} = 1$ kHz, $f_{z2} = 10$ kHz, $f_{p1} = 100$ kHz, $f_{p2} = 1$ MHz, and $|H_{EA}(j316\text{ kHz})| = 10$ dB.

11.38 The EA implementation with Z_A and Z_B as pictured in Fig. 11.37 is not unique. An alternative that you may find in the application notes uses Z_B to create f_{z2} and f_{p2}, and Z_A to create f_{z1} and f_{p1}. Use this alternative to design the EA of Problem 11.37.

11.8 Voltage mode control

11.39 With reference to the circuit of Fig. 11.40, derive the transfer function from the source at the left to the output at the right, assuming negligible loading by R_{LD} and Z_A. Hence, prove Eq. (11.52).

11.40 Redesign the EA of Example 11.22, but using the alternative implementation of Problem 11.38.

11.41 The op amp of Example 11.22 was assumed to have GBP $= \infty$ for simplicity. If GBP is noninfinite we expect some phase-margin erosion. Specify GBP_{min} for a maximum erosion of about 5°, and verify with PSpice.

11.42 Estimate the change in the average value V_O brought about by a 1-A change in the average value I_O in the converter of Fig. 11.42. Verify with PSpice.

11.9 Peak current mode control

11.43 Show that if $S_e = -S_f/2$, the Q factor of Eqs. (11.61) and (11.70) simplifies as $Q = 2/[\pi(1 - D)]$.

11.44 Reconsider Example 11.24, but for the case in which 1/2 of the output current is drawn by I_{LD}, and the other half by R_{LD}. Which parameters change, and how? What is the phase margin with the modified load?

11.45 A beginner, encouraged by the fact that PCMC eliminates the coil delay from the loop, tries to use just an ordinary op amp as the error amplifier and no additional components to compensate the converter of Example 11.24. (*a*) If the op amp has GBP = 1 MHz, will the circuit work? Explain! (*b*) Assuming the GBP is adjustable, find the GBP required for a phase margin of 45°. Verify with PSpice, compare with Example 11.24, and comment.

11.46 Design an EA to stabilize the converter of Example 11.24, but using an op amp instead of an OTA. What is a reasonable minimum GBP for the op amp to have? Verify with PSpice.

11.47 A buck converter with $V_I = 12$ V and $V_O = 5$ V is terminated on a 1-Ω load. Assuming $L = 10\,\mu$H, $C = 50\,\mu$F, $f_S = 350$ kHz, $V_p = 0.2$ V, $R_{SW} = R_{\text{sense}} = 20$ mΩ, $a_i = 5$ V/V, $R_{\text{coil}} = R_{\text{cap}} = 5$ mΩ, and $V_{\text{REF}} = 1.25$ V, design a compensator using an op amp.

11.10 PCMC of boost converters

11.48 (*a*) Show that the circuit of Fig. 11.60*a* has an LHP pole at $s = -1/(RC)$ and an RHP zero at $s = +R_1/(R_2RC)$. (*b*) If $R_1 = R_2$, the circuit is the familiar phase shifter of Fig. 3.12. Show its step response. (*c*) Repeat if $R_2 = 2R_1$, comment.

11.49 Design an EA to stabilize the converter of Example 11.25, but using an op amp instead of an OTA. What is a reasonable minimum GBP for the op amp to have? Verify with PSpice.

REFERENCES

1. Analog Devices Engineering Staff, *Practical Design Techniques for Power and Thermal Management,* Analog Devices, Norwood, MA, 1988. ISBN-0-916550-19-2.
2. P. R. Gray, P. J. Hurst, S. H. Lewis, and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits,* 5th ed., John Wiley & Sons, New York, 2009. ISBN 978-0-470-24599-6.
3. R. Knapp, "Selection Criteria Assist in Choice of Optimum Reference," *EDN,* February 18, 1988, pp. 183–192.
4. "IC Voltage Reference Has 1 ppm per Degree Drift," Texas Instruments Application Note AN-161, http://www.ti.com/lit/an/snoa589b/snoa589b.pdf.
5. A. P. Brokaw, "A Simple Three-Terminal IC Bandgap Reference," *IEEE Journal of Solid-State Circuits,* Vol. SC-9, No. 6, December 1974, pp. 388–393.
6. Analog Devices Engineering Staff, *Practical Design Techniques for Sensor Signal Conditioning,* Analog Devices, Norwood, MA, 1999. ISBN-0-916550-20-6.
7. J. Graeme, "Precision DC Current Sources," Part 1 and Part 2, *Electronic Design,* Apr. 26, 1990, pp. 191–198 and 201–206.
8. "Applications for an Adjustable IC Power Regulator," Texas Instruments Application Note AN-178, http://www.ti.com/lit/an/snva513a/snva513a.pdf.
9. G. A. Rincon Mora and P. E. Allen, "Study and Design of Low Drop-Out Regulators," School of Electrical and Computer Engineering, Georgia Institute of Technology, http://users.ece.gatech.edu/~rincon/publicat/journals/unpub/ldo_des.pdf.
10. R. W. Erikson and D. Maksimovic, *Fundamentals of Power Electronics,* 2nd ed., Springer Science + Business Media, Dardrecht, Netherlands, 2001. ISBN 00-052569.
11. C. Nelson, "LT1070 Design Manual," Linear Technology Application Note AN-19, http://cds.linear.com/docs/en/application-note/an19fc.pdf.
12. C. Nelson, "LT1074/LT1076 Design Manual," Linear Technology Application Note AN-44, http://cds.linear.com/docs/en/application-note/an44fa.pdf.
13. R. Ridley, "A New, Continuous-Time Model for Current-Mode Control," *IEEE Trans. on Power Electronics*, Vol. 6, No. 2, April 1991, pp. 272–280.
14. R. Ridley, "Current-Mode Control Modeling," *Switching Power Magazine,* 2006, pp. 1–9, http://www.switchingpowermagazine.com/downloads/5%20Current%20Mode%20Control%20Modeling.pdf.
15. R. Sheenan, "Understanding and Applying Current-Mode Control Theory," Texas Instruments Document SNVA555, 2007, http://www.ti.com/lit/an/snva555/snva555.pdf.
16. T. Hegarty, "Peak Current-Mode DC-DC Converter Stability Analysis," 2010, http://powerelectronics.com/regulators/peak-current-mode-dc-dc-converter-stability-analysis.

12

D-A AND A-D CONVERTERS

传递信息的变量（例如电压、电流、电荷、温度、压力等）在自然状态下是以模拟形式存在。然而，出于处理、传递和存储的目的，通常用数字形式表示信息更为方便。举例而言，考虑一个运算放大器，它需要输出一个电压范围为 0 ~ 1V 的信号 v，该信号的精度要求是 1mV 或者 0.1%。如果考虑器件的非理想性、漂移、老化、噪声、非完美的接线和互联，那么甚至一般的精度要求都很难满足。

如果信息用数字形式表示，那么对于电路性能的要求将大大放宽。举例而言，采用人们最常用的十进制形式，上述信号能表示为 $v=0.d1d2\cdots d_n$，其中 $d1$，$d2$，$\cdots$，d_n 是 0 ~ 9 之间的十进制数字量。在 $0.000\text{V}\le v<0.999\text{V}$ 的范围内，需要用 3 位这种数字量来表示 1mV 的分辨率。这就需要 3 个独立的电路来保持每个单独的数字量。每个电路只需要分辨 10 个电压等级而不是 1000 个电压等级，因此对于电路性能的要求就降低了。完成这个任务只需要每个电路具有 ± 5% 的精度。

信号数字化虽然解决了对于电路性能要求的问题，但是随即产生了需要进

行模数（A-D）和数模（D-A）转换的问题。举例而言，在上例中相应的电路需要为一个十进制的 D-A 转换器提供 $d1$、$d2$ 和 $d3$ 的值（这很容易实现），同时，需要用 1mV 精度合成 $v=d_1 10^{-1}+d_2 10^{-2}+d_3 10^{-3}$ 的模拟信号（这很难实现）。

尽管对于人们来说很方便，但是十进制表示并没有最大限度地减小对于电路性能的要求。仅用 0 和 1 两个值却能做得到这一点。如果我们用差异足够大的电压（例如 0V 和 5V）来表示这两个值，那么即使最粗糙的电路都能够分辨它们。因此，数字二进制或者位精确地构成了数字系统的基础。二进制电路（例如开关、逻辑门、触发器）可以用来保持和处理位。

图 12.1 描绘了 A-D 和 D-A 转换器最典型的应用背景 [1]。在适当处理后，一个模拟输入信号经过 A-D 转换器，用数字信号处理器（DSP）模块进行处理，或许只是以数字形式进行传输或记录。经过一次处理、接收或者恢复后，用 D-A 转换器将信号转换成模拟形式，也许再加一些额外的输出条件。

A-D 转换器（ADC）每秒采样 f_S 次。为了避免任何混叠现象的出现 [1]，模拟输入信号的最大频率分量必须被限制在 $f_S/2$ 以下；抗混叠滤波器参见第 4 章。ADC 通常需要在转换过程中输入信号保持不变，这表示 ADC 在每次转换之前需要一个采样保持电路来冻结带宽受限的信号；采样保持电路参见第 9 章。D-A 转换器（DAC）通常工作在和 ADC 相同的速率 f_S，如果应用需要，它通常需要具备一些合适的电路来消除在输入编码变化时输出出现的毛刺。阶梯状结果最后通过一个平滑滤波器来缓解量化噪声的影响。

图 12.1 的电路被无数次地以整体或者部分的形式应用。例如，数字信号处理（DSP）、直接数字控制（DDC）、数字音频混合、记录和回放、脉冲编码调制（PCM）通信、数据采集、电脑音乐、视频合成和数字万用表等。

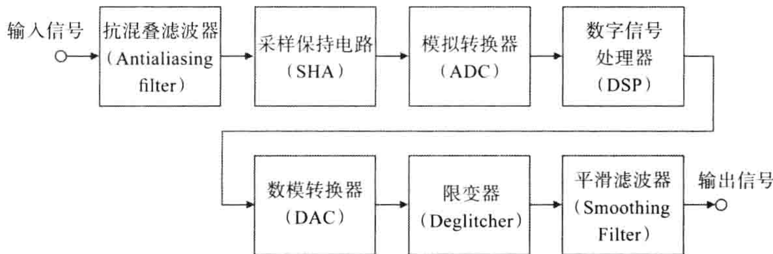


图 12.1　数据采样系统

本章重点

在介绍了 A-D 与 D-A 一些重要的专用名词和性能参数后，本章讨论常用的

D-A 转换技术，例如电阻权重和电容权重 DAC、电位 DAC、电流模式和电压模式 R-2R 阶梯 DAC、主从 DAC、电压模式和电流模式分段 DAC。一个典型的 R-$2R$ 阶梯 DAC，称为乘法 DAC，在许多应用中出现，最常见的包括数字可编程衰减器、滤波器和振荡器。

本章也介绍常用的 A-D 转换技术。首先讨论基于 DAC 的 ADC，例如逐次逼近和电荷分配 ADC。其次介绍高速转换器、辅助量程转换器和流水线转换器。最后讨论积分形式的 ADC，例如电荷平衡和双斜率转换器。

本章以过采样转换器结尾，并讨论与之有关的奈奎斯特采样、过采样、噪声整形和 sigma-delta 转换器。

12.1 PERFORMANCE SPECIFICATIONS

A string of n bits, $b_1b_2b_3\dots b_n$, forms an n-bit word. Bit b_1 is called the *most significant bit* (MSB) and bit b_n the *least significant bit* (LSB). The quantity

$$D = b_1 2^{-1} + b_2 2^{-2} + b_3 2^{-3} + \cdots + b_n 2^{-n} \tag{12.1}$$

is called the *fractional binary value*. Depending on the bit pattern, D can assume 2^n equally spaced values from 0 to $1 - 2^{-n}$. The lower limit is reached when all bits are 0, the upper limit when all bits are 1, and the spacing between adjacent values is 2^{-n}.

D-A Converters (DACs)

A DAC accepts an n-bit input word $b_1b_2\dots b_n$ with fractional binary value D_I, and produces an analog output proportional to D_I. Figure 12.2*a* depicts a voltage-output DAC, for which we have

$$v_O = KV_{\text{REF}}D_I = V_{\text{FSR}}(b_1 2^{-1} + b_2 2^{-2} + \cdots + b_n 2^{-n}) \tag{12.2}$$

where K is a *scale factor;* V_{REF} is a *reference voltage;* b_k $(k = 1, 2, \dots, n)$ is either 0 or 1, depending on the logic level at the corresponding input; and $V_{\text{FSR}} = KV_{\text{REF}}$ is the *full-scale range.* Frequently used values for V_{FSR} are 2.5 V, 5.0 V, and 10.0 V. Though our discussion will focus on voltage-output DACs, the results are readily extended to current-output DACs, characterized by $i_O = KI_{\text{REF}}D_I = I_{\text{FSR}}D_I$. A typical I_{FSR} value is 1.0 mA.

We observe that the DAC output is the result of multiplying the analog signal V_{REF} by the digital variable D_I. A DAC that allows for V_{REF} to vary all the way down to zero is called a *multiplying DAC* (MDAC).

Depending on the input bit pattern, v_O can assume 2^n different values ranging from 0 to the *full-scale value* $V_{\text{FSV}} = (1 - 2^{-n})V_{\text{FSR}}$. The MSB contribution to v_O is $V_{\text{FSR}}/2$, and the LSB contribution is $V_{\text{FSR}}/2^n$. The latter is called the *resolution,* or simply the *LSB.* Note that V_{FSV} is always 1 LSB short of V_{FSR}. The quantity $\text{DR} = 20\log_{10} 2^n$ is called the *dynamic range* of the DAC. Thus,

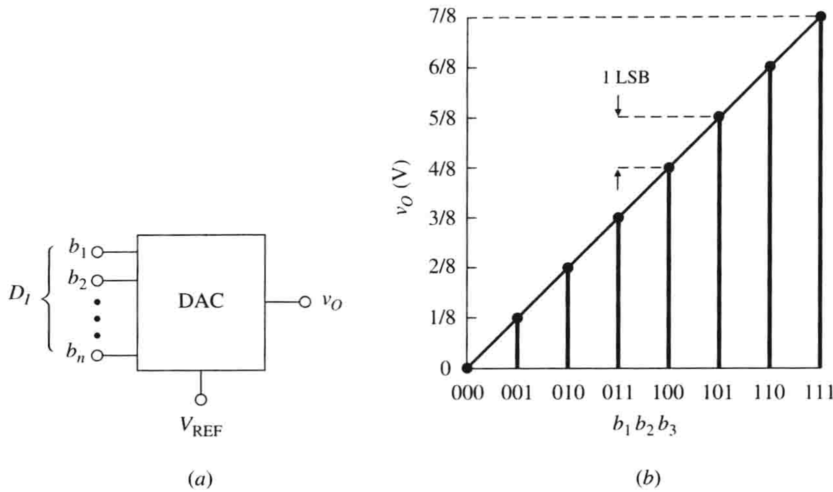


FIGURE 12.2
DAC diagram and ideal transfer characteristic for $n = 3$ and $V_{FSR} = 1$ V.

a 12-bit DAC with $V_{FSR} = 10.000$ V has LSB $= 2.44$ mV, $V_{FSV} = 9.9976$ V, and DR $= 72.25$ dB.

Since there are only 2^n possible input codes, the transfer characteristic of a DAC is a set of points whose envelope is a straight line with end points at $(b_1 b_2 \dots b_n, v_O) = (00 \dots 0, 0\text{ V})$ and $(11 \dots 1, V_{FSV})$. Figure 12.2*b* shows the characteristic of a DAC with $n = 3$ and $V_{FSR} = 1.0$ V. The graph consists of $2^3 = 8$ bars ranging in height from 0 to $V_{FSV} = \frac{7}{8}$ V with a resolution of 1 LSB $= \frac{1}{8}$ V. If we drive a DAC with a uniformly clocked n-bit binary counter and observe v_O with the oscilloscope, the waveform will be a staircase. The higher n, the finer the resolution and the closer the staircase to a continuous ramp. DACs are available in word lengths ranging from 6 bits to 20 bits or more. While DACs with 6, 8, 10, 12, and 14 bits are common and economical, DACs with $n > 14$ become progressively more expensive and require the utmost care to realize their full precision.

DAC Specifications[3]

The internal circuitry of a DAC is subject to component mismatches, drift, aging, noise, and other sources of error, whose effect is to degrade conversion performance. The maximum deviation of the actual output from the ideal value predicted by Eq. (12.2) is called the *absolute accuracy* and is expressed in fractions of 1 LSB. Clearly, if an n-bit DAC is to retain its credibility down to its LSB, its absolute accuracy must never be worse than $\frac{1}{2}$ LSB. DAC errors are classified as *static* and *dynamic*.

The simplest static errors are the *offset* error and the *gain* error depicted in Fig. 12.3. The offset error (+1 LSB in the example) is nulled by translating the actual envelope up or down until it goes through the origin, as in Fig. 12.3*b*. What is left, then, is the gain error (−2 LSB in the example), which is nulled by adjusting the scale factor K.

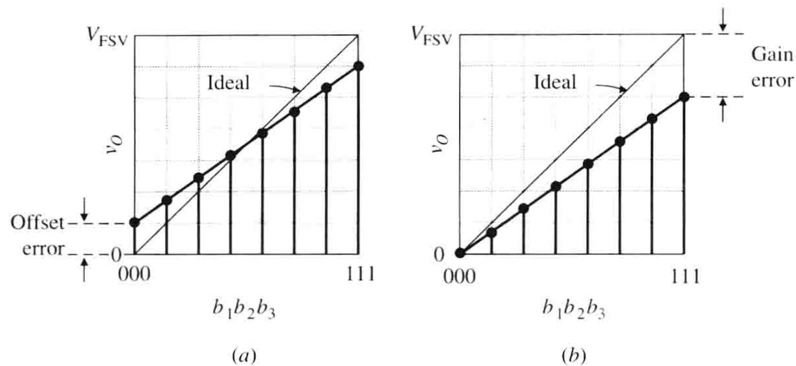


FIGURE 12.3
DAC offset error and gain error.

Even after both errors have been nulled, the actual envelope is likely to deviate from the straight line passing through the end points. The maximum deviation is called the *integral nonlinearity* (INL), or also the *relative accuracy,* and is expressed in fractions of 1 LSB. Ideally, the difference in height between adjacent bars is 1 LSB; the maximum deviation from this ideal value is called the *differential nonlinearity* (DNL). If DNL < -1 LSB, the transfer characteristic becomes *nonmonotonic;* that is, for certain input code transitions v_O will decrease with the input code, rather than increase. A nonmonotonic characteristic is especially undesirable in control, where it may cause oscillations, and in successive-approximation ADCs, where it may lead to missing codes. An example will better clarify these concepts.

EXAMPLE 12.1. Find the INL and DNL of the 3-bit DAC of Fig. 12.4. Comment on your results.

Solution. By inspection, the individual-code integral and differential nonlinearities, in fractions of 1 LSB, are found to be

k:	000	001	010	011	100	101	110	111
INL_k:	0	0	−1/2	1/2	−1	1/2	−1/2	0
DNL_k:	0	0	−1/2	1	−3/2	3/2	−1	1/2

The maxima of INL_k and DNL_k are, respectively, INL $= 1$ LSB and DNL $= 1\frac{1}{2}$ LSB. We observe a nonmonotonicity as the code changes from 011 to 100, where the step size is $-\frac{1}{2}$ LSB instead of $+1$ LSB; hence, $\text{DNL}_{100} = -\frac{1}{2} - (+1) = -\frac{3}{2}$ LSB $<$ -1 LSB. The fact that $\text{DNL}_{101} = \frac{3}{2}$ LSB $>$ 1 LSB, though undesirable, does not cause nonmonotonicity.

Remark. Note that $\text{INL}_k = \sum_{i=0}^{k} \text{DNL}_i$. Can you provide an intuitive justification?

DAC performance changes with temperature, age, and power-supply variations; hence, all relevant performance parameters such as offset, gain, INL and DNL, and monotonicity must be specified over the full temperature and power-supply ranges.

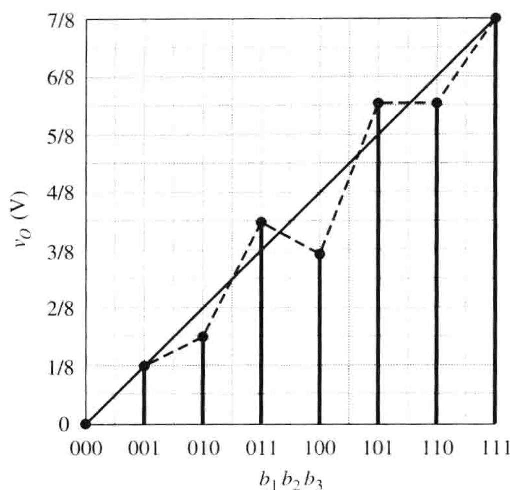


FIGURE 12.4
Example of actual DAC characteristics after the offset and gain errors have been nulled.

The most important dynamic parameter is the *settling time* t_S. This is the time it takes for the output to settle within a specified band (usually $\pm\frac{1}{2}$ LSB) of its final value following a code change at the input (usually a full-scale change). Typically, t_S ranges from under 10 ns to over 10 μs, depending on word length as well as circuit architecture and technology.

Another potential source of concern is the presence of output spikes in connection with major input-code transitions. Called *glitches,* these spikes are due to the internal circuitry's nonuniform response to input bit changes as well as poor synchronization of the bit changes themselves. For instance, if during the center-scale transition from 01 . . . 1 to 10 . . . 0 the MSB is perceived as going on before (or after) all other bits go off, the output will momentarily swing to full scale (or to zero), causing a positive-going (or negative-going) output spike, or glitch.

Glitches are of particular concern in CRT display applications. They can be minimized by synchronizing the input bit changes with a high-speed parallel latch register, or by processing the DAC output with a THA. The THA is switched to the hold mode just prior to the input code change, and is returned to the track mode only after the DAC has recovered from the glitch and settled to its new level.

A-D Converters (ADCs)

An ADC provides the inverse function of a DAC. As shown in Fig. 12.5*a*, it accepts an analog input v_I and produces an output word $b_1 b_2 \dots b_n$ of fractional value D_O such that

$$D_O = b_1 2^{-1} + b_2 2^{-2} + \cdots + b_n 2^{-n} = \frac{v_I}{K V_{\text{REF}}} = \frac{v_I}{V_{\text{FSR}}} \tag{12.3}$$

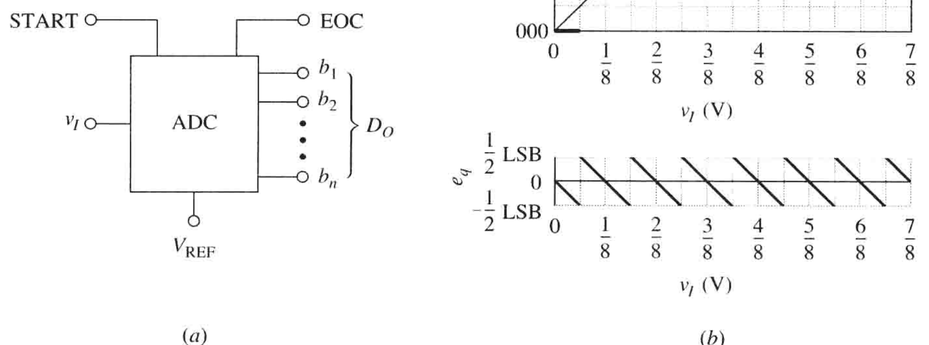


FIGURE 12.5
ADC diagram, and ideal transfer characteristic and quantization noise for $n = 3$ and $V_{FSR} = 1$ V.

Usually, an ADC includes two additional control pins: the START input, to tell the ADC when to start converting, and the EOC output, to announce when conversion is complete. The output code can be in either parallel or serial form. ADCs are often equipped with latches, control logic, and tristate buffers to facilitate microprocessor interfacing. ADCs intended for digital panel-meter applications are designed to drive LCD or LED displays directly.

The input to an ADC is often a transducer signal proportional to the transducer supply voltage V_S, or $v_I = \alpha V_S$ (a load cell is a typical example). In these cases it is convenient to use V_S also as the reference to the ADC, for then Eq. (12.3) simplifies as $D_O = \alpha V_S / K V_S = \alpha / K$, indicating a reference-independent conversion. Called *ratiometric conversion,* this technique allows for highly accurate conversions using references of only modest quality.

Figure 12.5*b*, top, shows the ideal characteristic of a 3-bit ADC with $V_{FSR} =$ 1.0 V. The conversion process partitions the analog input range into 2^n intervals called *code ranges,* and all values of v_I within a given code range are represented by the same code, namely, that corresponding to the midrange value. For example, code 011, corresponding to the midrange value $v_I = \frac{3}{8}$ V, actually represents all inputs within the range $\frac{3}{8} \pm \frac{1}{16}$ V. Due to the inability by the ADC to distinguish among different values within this range, the output code can be in error by as much as $\pm\frac{1}{2}$ LSB. This uncertainty, called *quantization error,* or also *quantization noise* e_q, is an inherent limitation of any digitization process. An obvious way to improve it is by increasing n.

As shown in Fig. 12.5*b*, bottom, e_q is a sawtooth-like variable with a peak value of $\frac{1}{2}$ LSB $= V_{FSR}/2^{n+1}$. Its rms value is readily found to be $E_q = (\frac{1}{2}\text{ LSB})/\sqrt{3}$ or

$$E_q = \frac{V_{FSR}}{2^n\sqrt{12}} \tag{12.4}$$

If v_I is a sinusoidal signal, the signal-to-noise ratio is maximized when v_I has a peak amplitude of $V_{FSR}/2$, or an rms value of $(V_{FSR}/2)/\sqrt{2}$. Thus, $\text{SNR}_{max} = 20\log_{10}[(V_{FSR}/2\sqrt{2})/(V_{FSR}/2^n\sqrt{12})]$, or

$$\text{SNR}_{max} = 6.02n + 1.76 \text{ dB} \tag{12.5}$$

Increasing n by 1 cuts E_q in half and increases SNR_{max} by 6.02 dB.

ADC Specifications[3]

Similar to the case of DACs, ADC performance is characterized in terms of *offset* and *gain errors, differential* and *integral nonlinearity,* and *stability.* However, ADC errors are defined in terms of the values of v_I at which *code transitions* occur. Ideally, these transitions occur at odd multiples of $\frac{1}{2}$ LSB, as shown in Fig. 12.5*b*. In particular, the first transition (000 $\rightarrow$ 001) occurs at $v_I = \frac{1}{2}$ LSB $= \frac{1}{16}$ V, and the last (110 $\rightarrow$ 111) at $v_I = V_{FSV} - \frac{1}{2}$ LSB $= V_{FSR} - \frac{3}{2}$ LSB $= \frac{13}{16}$ V.

The *offset error* is the difference between the actual location of the first code transition and $\frac{1}{2}$ LSB, and the *gain error* is the difference between the actual locations of the last and first transition, and the ideal separation of $V_{FSR} - 2$ LSB. Even after both errors have been nulled, the locations of the remaining code transitions are likely to deviate from their ideal values, as exemplified in Fig. 12.6.

The dotted curve, representing the locus of the midpoints of the actual code ranges, is called the *code center line*. Its maximum deviation from the straight line

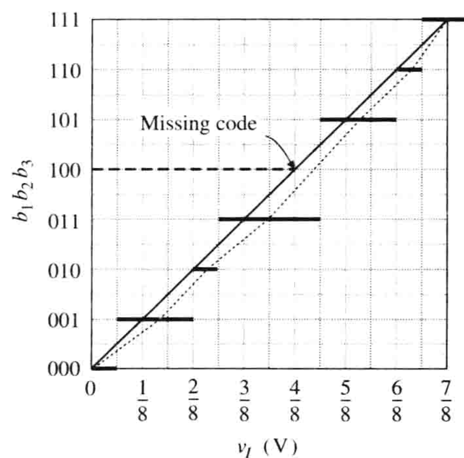


FIGURE 12.6
Example of actual ADC characteristic with missing code.

passing through the end points after the offset and gain errors have been nulled is called the *integral nonlinearity* (INL). Ideally, code transitions are 1 LSB apart. The maximum deviation from this ideal value is called the *differential nonlinearity* (DNL). If the DNL exceeds 1 LSB, some codes may be skipped at the output. Missing codes are undesirable in digital control, where they may lead to instability.

In the example shown, the INL error is maximized in connection with the 011 code range, where this error is $-\frac{1}{2}$ LSB. This range also maximizes the DNL error. The range width of 2 LSB indicates that DNL $=(2-1)$ LSB $=1$ LSB. Not suprisingly, there is a missing code. As you investigate INL and DNL errors, make sure you measure them along the horizontal (or the vertical) axis, not as geometric distances! As a check, you can use the relationship $\mathrm{INL}_k = \sum_{i=0}^{k} \mathrm{DNL}_i$, which holds also for ADCs.

An A-D conversion takes a certain amount of time to complete. Called the *conversion time,* it typically ranges from less than 10 ns to tens of milliseconds, depending on the conversion method, resolution, and technology.

A practical ADC will produce noise in excess of the theoretical quantization noise of Eq. (12.4). It will also introduce distortion due to transfer-characteristic nonlinearities. The effective number of bits is then[4]

$$\mathrm{ENOB} = \frac{S/(N+D) - 1.76\ \mathrm{dB}}{6.02} \tag{12.6}$$

where $S/(N+D)$ is the actual signal-to-noise-plus-distortion ratio, in decibels.

EXAMPLE 12.2. A 10-bit ADC with $V_{FSR} = 10.24$ V is found to have $S/(N+D) =$ 56 dB. Find E_q, SNR_{max}, and ENOB.

Solution. Using Eqs. (12.4) through (12.6) gives $E_q = 2.89$ mV, $\mathrm{SNR}_{max} = 61.97$ dB, and ENOB $= 9.01$, indicating nine effective bits. In other words, the given 10-bit ADC yields the same performance as an ideal 9-bit ADC.

12.2 D-A CONVERSION TECHNIQUES

DACs are available in a variety of architectures and technologies.[3,4] In this section we examine the most common examples.

Weighted-Resistor DACs

Equation (12.2) indicates that the functions required to implement an n-bit DAC are n switches and n binary-weighted variables to synthesize the terms $b_k 2^{-k}$, $k = 1, 2, \ldots, n$; moreover, we need an n-input summer, and a reference. The DAC of Fig. 12.7 uses an op amp to sum n binary-weighted currents derived from V_{REF} via the current-scaling resistances $2R, 4R, 8R, \ldots, 2^n R$. Whether the current $i_k = V_{REF}/2^k R$ appears in the sum depends on whether the corresponding switch is closed ($b_k = 1$) or open ($b_k = 0$). Writing $v_O = -R_f i_O$ gives

$$v_O = (-R_f/R) V_{REF} (b_1 2^{-1} + b_2 2^{-2} + \cdots + b_n 2^{-n}) \tag{12.7}$$

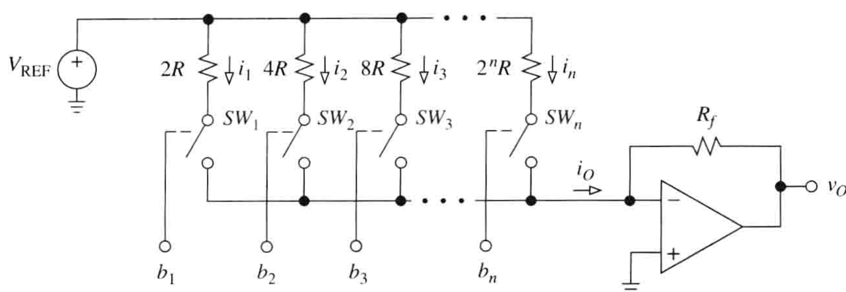


FIGURE 12.7
Weighted-resistor DAC.

indicating that $K = -R_f/R$. The offset error is nulled by trimming V_{OS}, and the gain error by adjusting R_f. Since the switches are of the virtual-ground type, they can be implemented with p-channel JFETs in the manner of Fig. 9.37.

The conceptual simplicity of the weighted-resistor DAC is offset by two drawbacks, namely, the nonzero resistances of the switches, and a spread in the current-setting resistances that increases exponentially with n. The effect of switch resistances is to disrupt the binary-weighted relationships of the currents, particularly in the most significant bit positions, where the current-setting resistances are smaller. These resistances can be made sufficiently large to swamp the switch resistances; however, this may result in unrealistically large resistances in the least significant positions. For instance, an 8-bit DAC requires resistances ranging from $2R$ to $256R$. The difficulty in ensuring accurate ratios over a range this wide, especially in monolithic form, restricts the practicality of resistor-weighted DACs below 6 bits.

Weighted-Capacitor DACs

Complex MOS ICs such as CODECS and microcomputers require on-chip data conversion capabilities using only MOSFETs and capacitors, which are the natural components of this technology. The DAC of Fig. 12.8 can be viewed as the

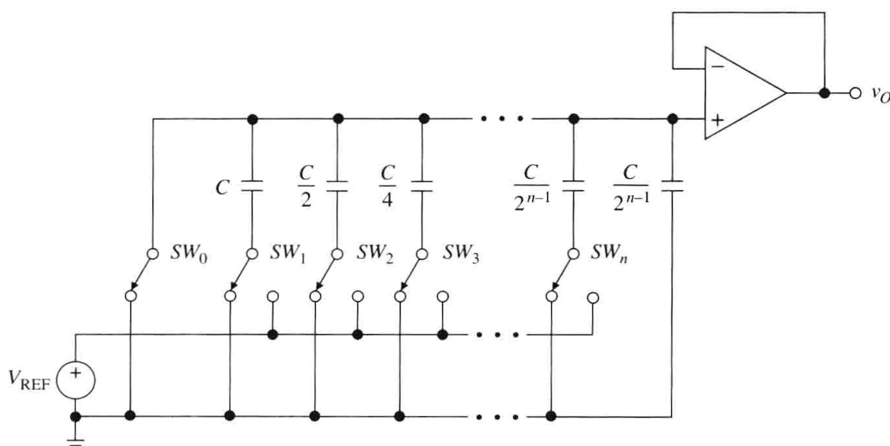


FIGURE 12.8
Weighted-capacitor DAC.

switched-capacitor counterpart of the weighted-resistor DAC just discussed. Its heart is an array of binary-weighted capacitances plus a terminating capacitance equal in value to the LSB capacitance. Circuit operation alternates between two cycles called the *reset* and *sample* cycles.

During the *reset cycle,* shown in the figure, all switches are connected to ground to completely discharge all capacitors. During the *sample cycle,* SW_0 is opened while each of the remaining switches is either left at ground or connected to V_{REF}, depending on whether the corresponding input bit is 0 or 1, respectively. This results in a redistribution of charge whose effect is to yield a code-dependent output.

Using elementary capacitor-divider principles, we readily find $v_O = V_{REF}\, C_r/C_t$, where C_r represents the sum of all capacitances connected to V_{REF}, and C_t the total capacitance of the array. We can write $C_r = b_1C + b_2C/2 + \cdots + b_nC/2^{n-1}$; moreover, $C_t = C + C/2 + \cdots + C/2^{n-1} + C/2^{n-1} = 2C$. Substituting gives

$$v_O = V_{REF}(b_1 2^{-1} + b_2 2^{-2} + \cdots + b_n 2^{-n}) \qquad (12.8)$$

indicating that the sample cycle provides an n-bit D-A conversion with $V_{FSR} = V_{REF}$.

By the artifice of switching the bottom plates, as shown, the bottom-plate parasitic capacitances are connected either to ground or to V_{REF}, without affecting charge distribution in the active capacitances. Since MOS capacitance ratios are easily controlled to 0.1% accuracies, the weighted-capacitor scheme is suitable for $n \leq 10$. As with weighted-resistor DACs, the main drawback of this scheme is an exponentially increasing capacitance spread.

Potentiometric DACs

It is not difficult to imagine the impact that component mismatches in the most significant bit positions of the previous DACs may have on differential nonlinearity and monotonicity. A *potentiometric* DAC achieves inherent monotonicity by using a string of 2^n resistors to partition V_{REF} into 2^n identical intervals. As depicted in Fig. 12.9 for $n = 3$, a binary tree of switches then selects the tap corresponding to the given input code and connects it to a high-input-impedance amplifier with gain $K = 1 + R_2/R_1$.

No matter how mismatched the resistors, v_O will always increase as the amplifier is switched from one tap to the next, up the ladder, hence the inherent monotonicity. Another advantage is that if the top and bottom nodes of the resistive string are biased at some arbitrary voltages V_H and V_L, the DAC will interpolate between V_L and V_H with a resolution of 2^n steps. However, the large number of resistors (2^n) and switches ($2^{n+1} - 2$) required limits practical potentiometric DACs to $n \leq 8$, even though the switches can be fabricated very efficiently in MOS technology.

R-2*R* Ladders

Most DAC architectures are based on the popular R-$2R$ ladder depicted in Fig. 12.10. Starting from the right and working toward the left, one can readily prove that the equivalent resistance to the right of each labeled node equals $2R$. Consequently, the

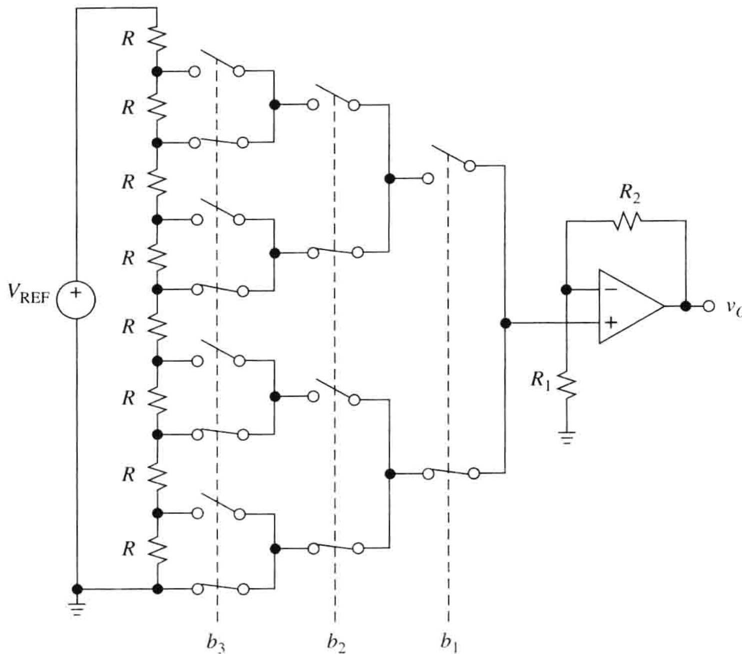


FIGURE 12.9
Potentiometric DAC.

current flowing downward, away from each node, is equal to the current flowing toward the right; moreover, twice this current enters the node from the left. The currents and, hence, the node voltages are binary-weighted,

$$i_{k+1} = \tfrac{1}{2} i_k \qquad v_{k+1} = \tfrac{1}{2} v_k \tag{12.9}$$

$k = 1, 2, \ldots, n - 1$. (Note that the rightmost $2R$ resistance serves a purely terminating function.)

With a resistance spread of only 2-to-1, R-$2R$ ladders can be fabricated monolithically to a high degree of accuracy and stability. Thin-film ladders, fabricated by deposition on the oxidized silicon surface, lend themselves to accurate laser trimming for DACs with $n \geq 12$. For DACS with a lower number of bits, diffused or

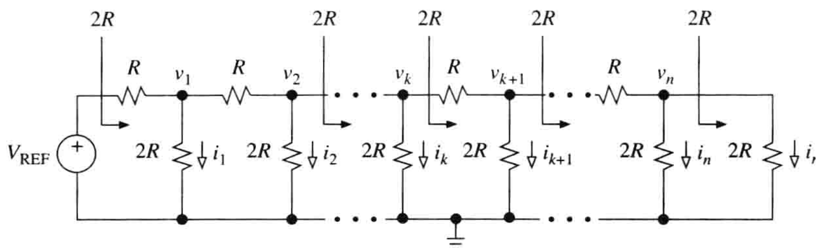


FIGURE 12.10
R-$2R$ ladder.

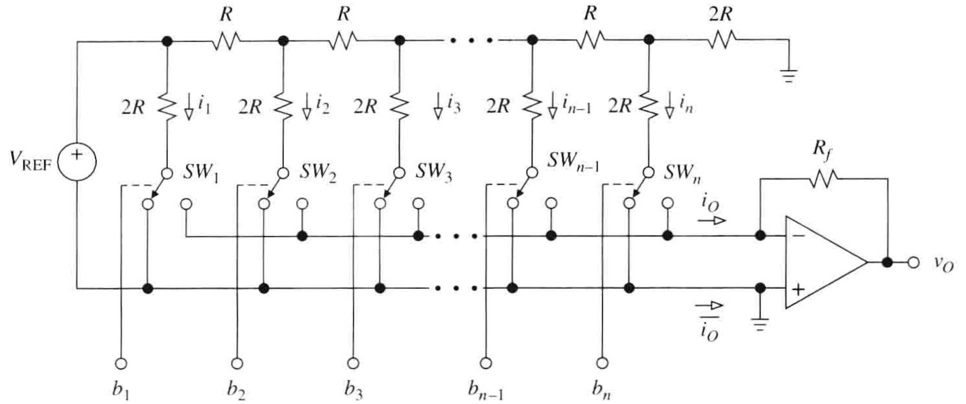


FIGURE 12.11
DAC using a current mode *R*-2*R* ladder.

ion-implanted ladders are often adequate. Depending on how the ladder is utilized, different DAC architectures result.

Current Mode *R*-2*R* Ladder

The architecture of Fig. 12.11 derives its name from the fact that it operates on the ladder currents. These currents are $i_1 = V_{REF}/2R = (V_{REF}/R)2^{-1}$, $i_2 = (V_{REF}/2)/2R = (V_{REF}/R)2^{-2}, \ldots, i_n = (V_{REF}/R)2^{-n}$, and they are diverted either to the ground bus ($\overline{i_O}$) or to the virtual-ground bus (i_O). Using bit b_k to identify the status of SW_k, and letting $v_O = -R_f i_O$ gives

$$v_O = -(R_f/R)V_{REF}(b_1 2^{-1} + b_2 2^{-2} + \cdots + b_n 2^{-n}) \tag{12.10}$$

indicating that $K = -R_f/R$. Since $i_O + \overline{i_O} = (1 - 2^{-n})V_{REF}/R$ regardless of the input code, $\overline{i_O}$ is said to be *complementary* to i_O. An important advantage of the current mode is that the voltage change across each switch is minimal, so charge injection is virtually eliminated and switch-driver design is made simpler.

We observe that the potential of the i_O bus must be sufficiently close to that of the $\overline{i_O}$ bus; otherwise, linearity errors will occur. Thus, in high-resolution DACs, it is crucial that the overall input offset error of the op amp be nulled and have low drift.

Voltage Mode *R*-2*R* Ladder

In the alternative mode of Fig. 12.12, the $2R$ resistances are switched between V_L and V_H, and the output is obtained from the leftmost ladder node. As the input code is sequenced through all possible states from $0\ldots0$ to $1\ldots1$, the voltage of this node changes in steps of $2^{-n}(V_H - V_L)$ from V_L to $V_H - 2^{-n}(V_H - V_L)$. Buffering it with an amplifier results in the scale factor $K = 1 + R_2/R_1$. The advantage of this scheme is that it allows us to interpolate between any two voltages, neither of which need necessarily be zero.

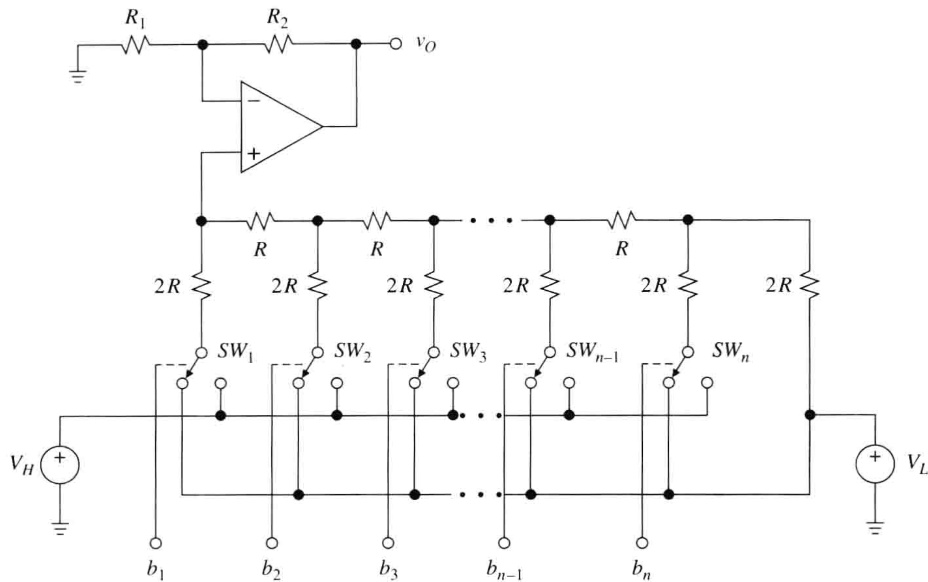


FIGURE 12.12
DAC using a voltage mode R-$2R$ ladder.

Bipolar DACs

In the architecture exemplified in Fig. 12.13 for $n = 4$, the R-$2R$ ladder is used to provide the current bias for n binary-weighted BJT current sinks; n nonsaturating BJT

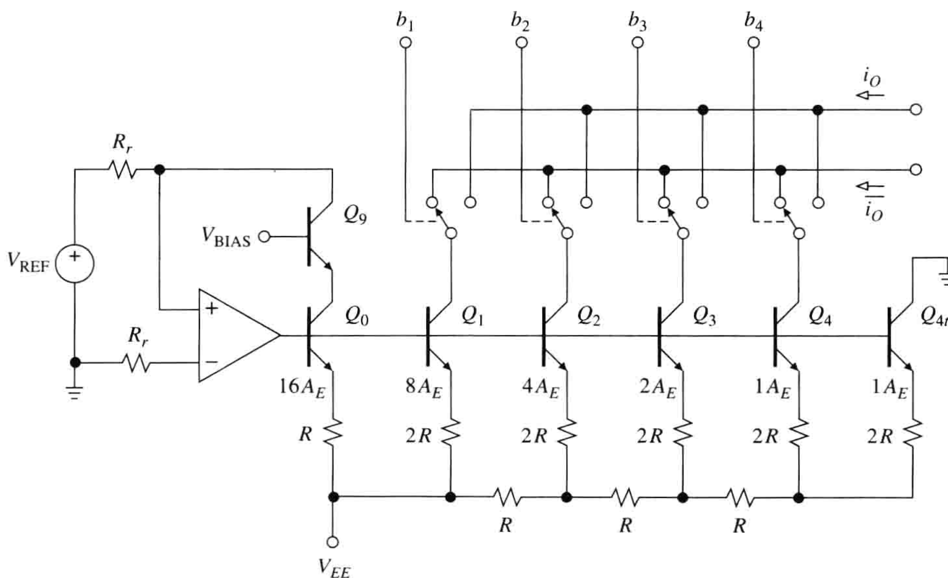
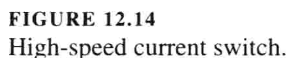


FIGURE 12.13
Bipolar DAC.



(a) (b)

FIGURE 12.14
High-speed current switch.

switches then provide fast current steering, typically in the range of nanoseconds. The current sinks are Q_1 through Q_4, with Q_{4t} providing a terminating function. We observe that for the ladder to work properly, the upper nodes of the $2R$ resistances must be equipotential. The voltages at these nodes are set by the emitters of the current sinks. Since the corresponding currents are in 2:1 ratios, the emitter areas must be scaled accordingly as $1A_E$, $2A_E$, $4A_E$, and $8A_E$ to ensure identical V_{BE} drops and, hence, equipotential emitters.

Figure 12.14 shows the details of the kth current-steering switch. For $v_k > V_{\text{BIAS1}}$, Q_1 is off and Q_2 is on. This, in turn, keeps Q_3 off and Q_4 on, thus steering the collector current of Q_k to the i_O bus. For $v_k < V_{\text{BIAS1}}$, the conditions are reversed and the current of Q_k is now diverted to the $\overline{i_O}$ bus. The switching threshold is typically set at $V_{\text{BIAS1}} \cong 1.4$ V to provide both TTL and CMOS compatibility.

We observe that because of the finite betas of the BJTs, the current losses in the bases introduce errors. The circuit of Fig. 12.13 uses Q_0 to compensate for the base losses of the current sinks, and Q_9 to compensate for the base losses of the switches. The circuit works as follows: by op amp action, $i_{C9} = V_{\text{REF}}/R_r$. Using the BJT relationship $i_C = \alpha i_E$, and assuming the same α throughout, we have $i_{E0} = i_{C0}/\alpha = i_{E9}/\alpha = (i_{C9}/\alpha)/\alpha = (V_{\text{REF}}/R_r)/\alpha^2$. By ladder action, the emitter current of the kth sink is $i_{Ek} = i_{E0}2^{-k}$. The kth current reaching the i_O bus is $i_k = \alpha i_{Ck} = \alpha(\alpha i_{Ek}) = \alpha^2 i_{E0}2^{-k} = (V_{\text{REF}}/R_r)2^{-k}$, indicating the disappearance of base current errors. Summing the various currents on the i_O bus gives

$$i_O = I_{\text{REF}}(b_1 2^{-1} + b_2 2^{-2} + b_3 2^{-3} + b_4 2^{-4}) \tag{12.11}$$

where $I_{\text{REF}} = V_{\text{REF}}/R_r$.

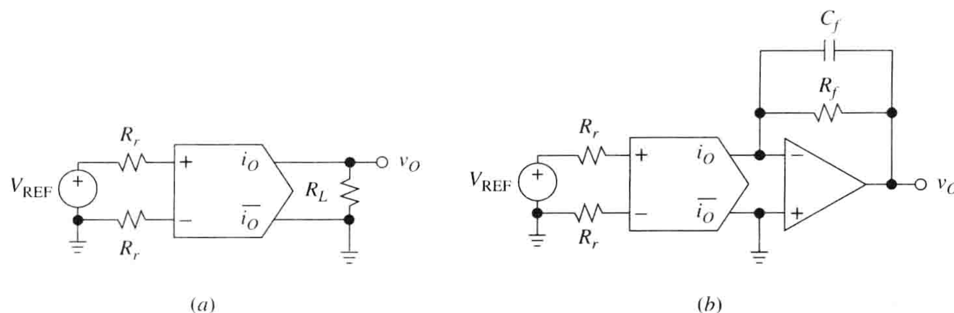


FIGURE 12.15
Bipolar DAC output conditioning.

Figure 12.15 shows the two most common ways of converting i_O to a voltage. The purely resistive termination of Fig. 12.15*a*, giving $v_O = -R_L i_O$, realizes the full-speed capability of the DAC as long as R_L is sufficiently small to render the effect of the stray output capacitance of the DAC negligible. The output swing is in this case limited by the voltage compliance of the DAC, as given in the data sheets. The op amp converter of Fig. 12.15*b* gives $v_O = R_f i_O$ with a low output impedance, but at the price of a degradation in dynamics as well as the extra cost of the op amp. The overall settling time t_S can be estimated from the individual settling times of the DAC and the op amp as

$$t_S = \sqrt{t^2_{S(\mathrm{DAC})} + t^2_{S(\mathrm{OA})}} \tag{12.12}$$

The purpose of C_f is to stabilize the op amp against the stray output capacitance of the DAC.[5] Suitable op amps for this application are either high-SR, fast-settling JFET-input types, or CFA types.

Master-Slave DACs

The resolution of the basic structure of Fig. 12.13 can, in principle, be increased by using additional current sinks; however, maintaining ratioed emitter areas soon leads to extravagant BJT geometries. The architecture of Fig. 12.16 eases the geometry requirements by combining two DACs of the type just discussed in a master-slave configuration in which the current of the terminating BJT Q_{4t} of the master DAC is used to bias the slave DAC. This current, representing 1 LSB of the master DAC, is partitioned by the slave DAC into four additional binary-weighted currents, with Q_{8t} now providing the required termination. The result is an 8-bit DAC with $I_{REF} = V_{REF}/R_r$ and a resolution of $I_{REF}/2^8$. Popular master-slave DACs are the DAC-08 (8-bit) and the DAC-10 (10-bit), both of which settle within $\pm\frac{1}{2}$ LSB in 85 ns (typical) and provide output voltage compliance down to -10 V.

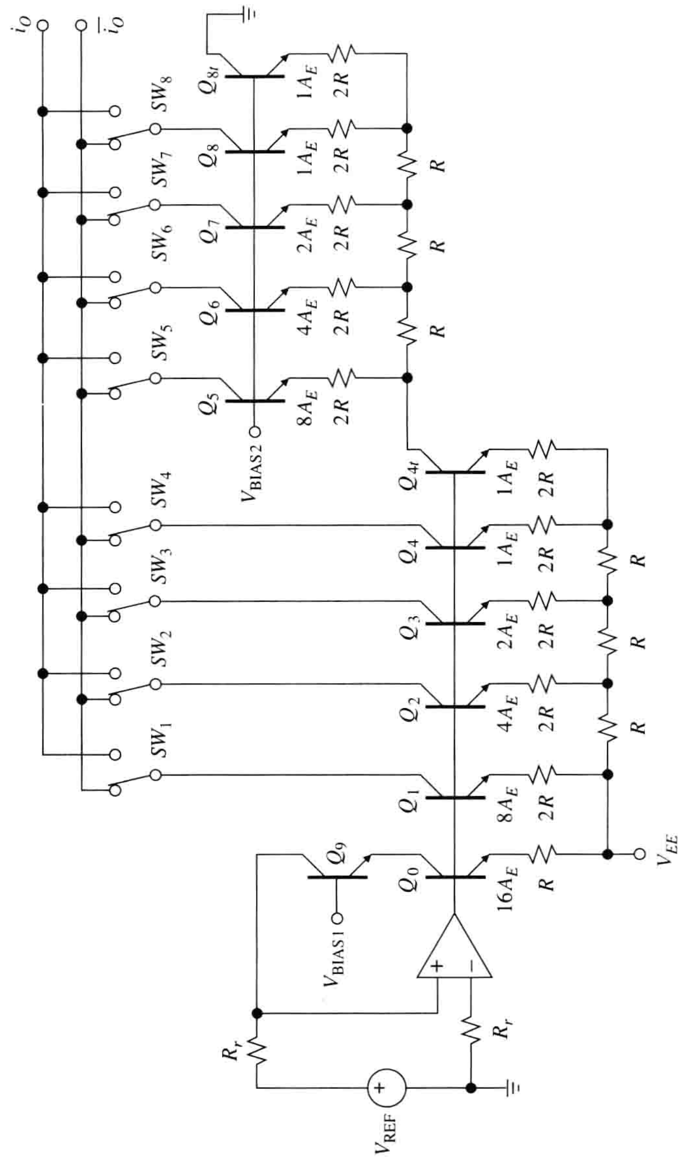


FIGURE 12.16
Master-slave DAC.

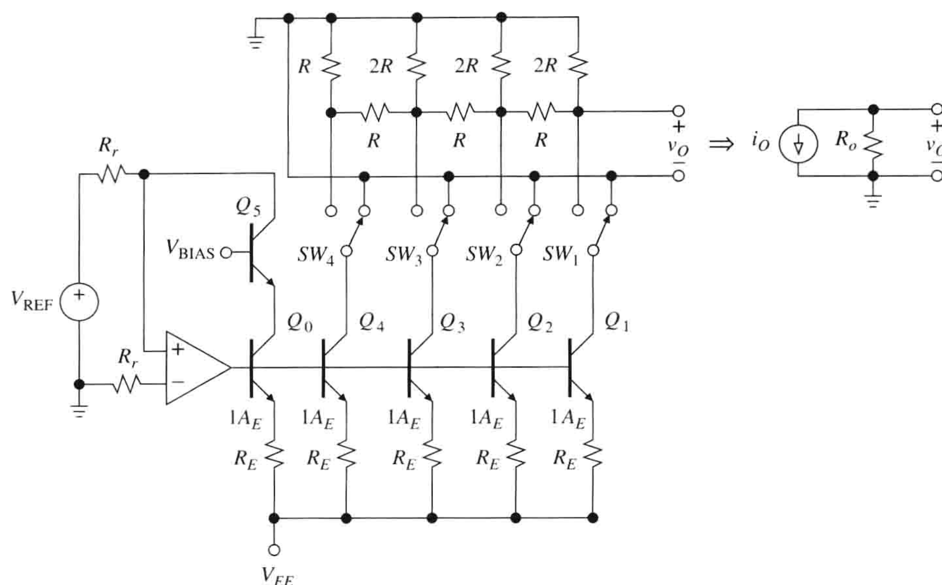


FIGURE 12.17
DAC using a current-driven R-$2R$ ladder.

Current-Driven *R*-2*R* Ladder

The problems stemming from emitter area scaling are eliminated altogether by using equal-value current sinks and exploiting the current-scaling capability of the R-$2R$ ladder to obtain binary-weighted contributions to the output. Though Fig. 12.17 shows a 4-bit example, the principle is readily extended to higher values of n. One can readily show (see Problem 12.8) that the ladder admits a Norton equivalent with $R_o = R$ and $i_O = (2V_{REF}/R_r)(b_1 2^{-1} + b_2 2^{-2} + b_3 2^{-3} + b_4 2^{-4})$; to reduce cluttering, b_1 through b_4 have been omitted.

The use of suitably small ladder resistances (≤ 1 kΩ) minimizes the effect of parasitic capacitances, allowing v_O to settle very rapidly. If the output is left floating, the DAC will give $v_O = -Ri_O = (-2R/R_r)V_{REF}D_I$ with $R_o = R$. Alternatively, if zero output impedance is desired, an I-V converter op amp can be used, but at the price of a longer settling time as per Eq. (12.12).

Segmentation

The matching and tracking capabilities of IC components limit the resolution of the DAC structures considered so far to $n \leq 12$. However, the areas of precision instrumentation and test equipment, process control, industrial weighing systems, and digital audio playback often require resolutions and linearity performance well in excess of 12 bits. One of the most important performance requirements is *monotonicity*. In fact, there are situations in which uniform step size in the DAC characteristics is more important than exact straight-line conformance. For instance, in process control, even though the inherent linearity of an input transducer may not surpass

0.1% or 10 bits, a higher number of bits is often required to resolve small transducer variations. Likewise, to ensure a high signal-to-noise ratio, digital audio playback systems use 16 bits or more of differential linearity, though not necessarily providing the same level of integral nonlinearity.

In conventional binary-weighted DACs, monotonicity is hardest to realize at the point of major carry due to the difficulty of realizing the required degree of match between the MSB and the combined sum of all remaining bits. To ensure monotonicity, this match must be better than one part in 2^{n-1}, indicating that difficulty increases exponentially with n. High-resolution DACs achieve monotonicity by a technique known as *segmentation.* Here the reference range is partitioned into a sufficiently large number of contiguous segments, and a DAC of lesser resolution is then used to interpolate between the extremes of the selected segment. We shall now discuss this technique for both voltage mode and current mode DACs.

Voltage Mode Segmentation

Figure 12.18 illustrates the segmentation technique utilized by the AD7846 16-bit DAC. The four MS input bits are decoded to select, via switches SW_0 through SW_{16}, one of sixteen voltage segments available along the resistor string. The selected segment is then buffered by the voltage followers and used as a reference voltage of nominal value $V_{REF}/16$ to drive a 12-bit voltage mode R-$2R$ DAC. The latter, in turn, partitions the selected segment into $2^{12} = 4096$ smaller steps, starting at the bottom of the segment and ending one step short of the top, to give

$$v_O = V_L + D_{12}(V_H - V_L) \qquad (12.13)$$

where V_H and V_L are, respectively, the top and the bottom of the selected segment, and D_{12} is the fractional value of the lower 12-bit code. Omitted from the figure for simplicity are an input latch register, the segment decoder and switch-driver circuitry, and an output deglitcher switch.

Since the 65,536 possible output levels consist of 16 groups of 4096 steps each, the major carry of the 12-bit DAC is repeated in each of the 16 segments. Consequently, the accuracy required of the string resistances to ensure a given differential nonlinearity is relaxed by a factor of 16. Note, however, that integral nonlinearity cannot be better than the accuracy of the string resistances. The AD7846 offers 16-bit monotonicity with an integral linearity error of ± 2 LSB, and a 9-μs settling time to 0.0003%.

Considering that with $V_{REF} = 10$ V the step size is only $10/2^{16} = 152\ \mu$V, op amp input offset errors could cause intolerable differential nonlinearity if the buffers were stepped up the ladder in fixed order. This problem is overcome by interchanging the buffers at each segment transition, a technique referred to as *leapfrogging.* This, in turn, requires that V_H and V_L also be interchanged to preserve the input polarity to the 12-bit DAC. This function is provided by SW_{OA_1} and SW_{OA_2}. The effect of buffer interchanging can be appreciated as follows.

With the switches positioned as shown, the DAC is processing segment 0. Denoting the input offset errors of the op amps as V_{OS1} and V_{OS2}, we have $V_H = V_1 + V_{OS1}$ and $V_L = 0 + V_{OS2}$, where $V_1 = V_{REF}/16$. The last level of segment 0 is found by inserting these expressions into Eq. (12.13) with $D_{12} = (1 - 2^{-12})$. This gives $v_{O(\text{last})} = V_1(1 - 2^{-12}) + V_{OS1} - (V_{OS1} - V_{OS2})2^{-12}$.

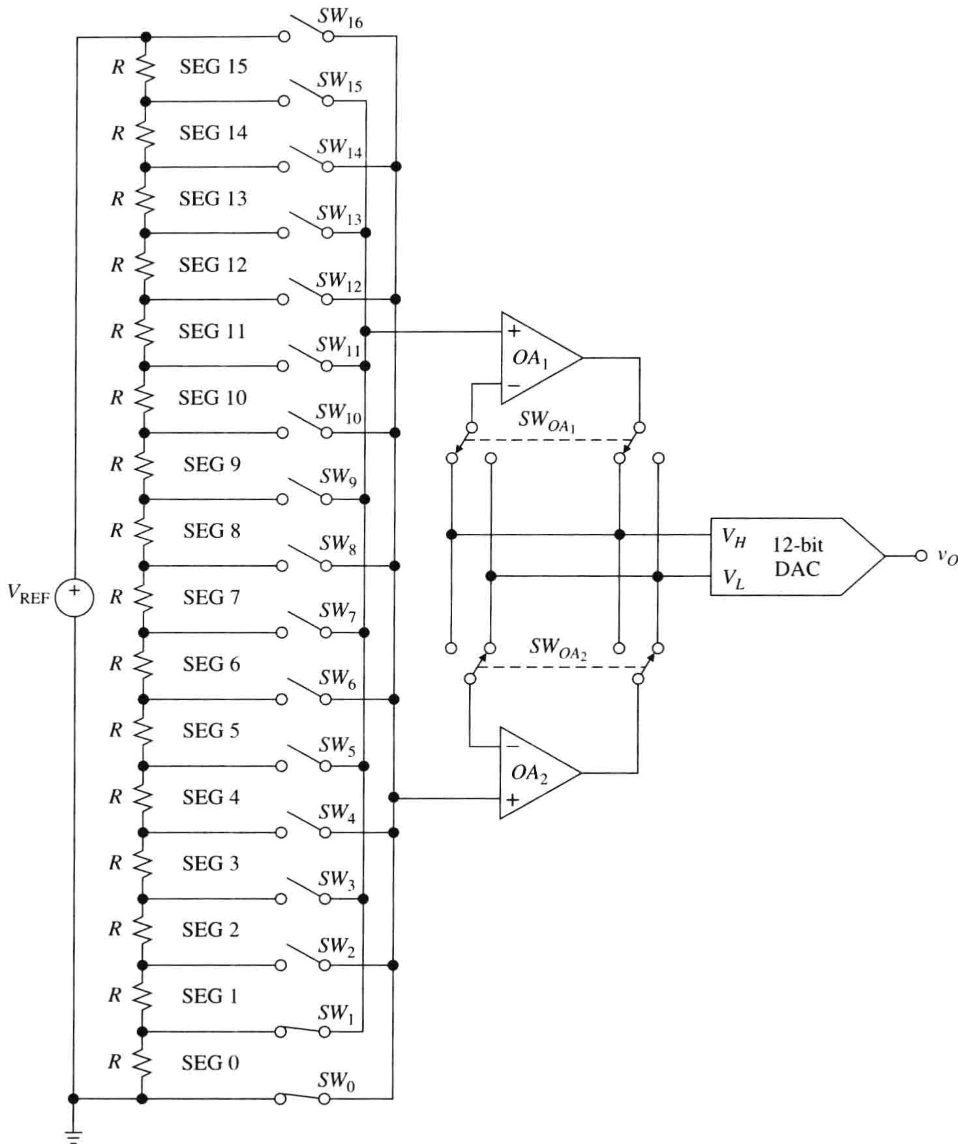


FIGURE 12.18
Simplified diagram of the AD7846 16-bit segmented DAC. (Courtesy of Analog Devices.)

At the point of transition from segment 0 to segment 1, SW_0 is opened, SW_1 and SW_2 are closed, and SW_{OA_1} and SW_{OA_2} are commutated. As a result, we now have $V_H = V_2 + V_{OS2}$ and $V_L = V_1 + V_{OS1}$, where $V_2 = 2V_1$. Consequently, the first level of segment 1 is $v_{O(\text{first})} = V_1 + V_{OS1}$. The difference between the two levels yields the step size at the first major carry,

$$v_{O(\text{first})} - v_{O(\text{last})} = \frac{V_{\text{REF}}}{2^{16}} + \frac{V_{OS2} - V_{OS1}}{2^{12}}$$

indicating that the leapfrogging technique reduces the combined offset error by 2^{12}. For instance, assuming $|V_{OS2} - V_{OS1}| \cong 10$ mV, the error term is $10^{-2}/2^{12} = 2.4\ \mu V \ll 1$ LSB. Similar considerations hold at the remaining segment transitions.

Current Mode Segmentation

Figure 12.19 illustrates segmentation for the case of a 16-bit current-mode R-$2R$ DAC. The resistances at the left establish 15 current segments of value V_{REF}/R, so the contribution of each segment to the output is $-(R_f/R)V_{\text{REF}}$. The decode logic examines the 4 MS input bits and diverts to the i_O bus 8 such segments for b_1, 4 segments for b_2, 2 segments for b_3, and 1 segment for b_4. The remaining resistances form an ordinary 12-bit current-mode R-$2R$ DAC, whose contribution to the output is given by Eq. (12.10). Using the superposition principle, we thus have $v_O = -(R_f/R)V_{\text{REF}} \times (8b_1 + 4b_2 + 2b_3 + b_4 + b_5 2^{-1} + b_6 2^{-2} + \cdots + b_{16} 2^{-12})$, or

$$v_O = -16\frac{R_f}{R}V_{\text{REF}}(b_1 2^{-1} + b_2 2^{-2} + \cdots + b_{16} 2^{-16}) \qquad (12.14)$$

indicating a 16-bit conversion with $V_{\text{FSR}} = -16(R_f/R)V_{\text{REF}}$. We observe that the segment resistances, like the ladder resistances, need only be accurate to 12 bits to ensure monotonicity at the 16-bit level. An example of a DAC using this principle is the MP7616 16-bit CMOS DAC.

Figure 12.20 shows a 16-bit segmented DAC using the current-driven ladder architecture. Here Q_1 through Q_7 provide 7 current segments of value $V_{\text{REF}}/4R_r = 0.25$ mA, which a decoder (not shown for simplicity) steers either to the i_O bus or to

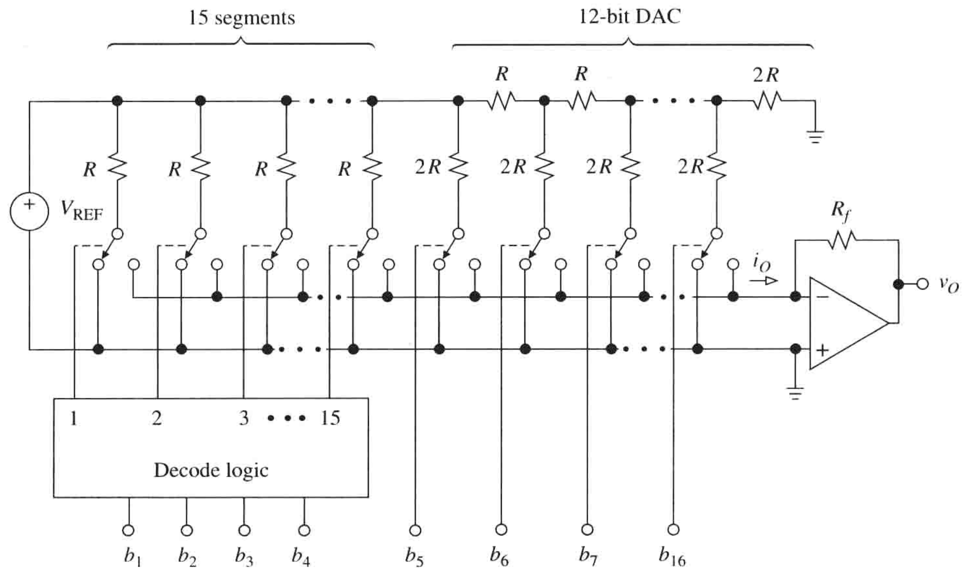


FIGURE 12.19
16-bit segmented DAC using a 12-bit current-mode R-$2R$ ladder.

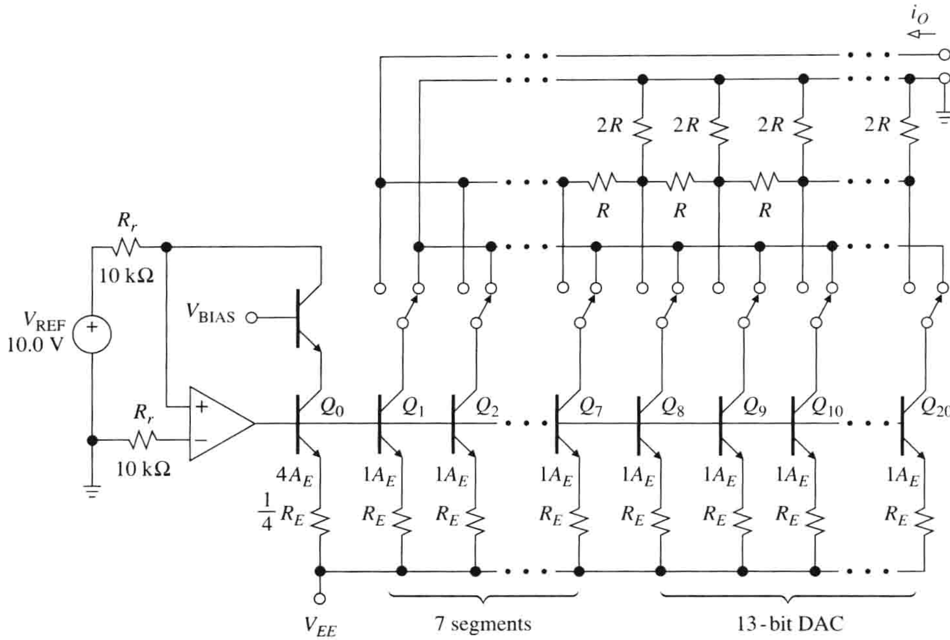


FIGURE 12.20
16-bit segmented DAC using a 13-bit current-driven R-$2R$ ladder.

ground, depending on the 3 MS bits. Steered to the i_O bus are 4 segments for b_1, 2 segments for b_2, and 1 segment for b_3. Moreover, Q_8 through Q_{20}, along with the R-$2R$ ladder, form a 13-bit current-driven DAC. Proper scaling requires an additional R resistance between the 13-bit DAC and the i_O bus. Consequently, the Norton resistance is now $R_o = 2R$. By the superposition principle, $i_O = (V_{REF}/4R_r)(4b_1 + 2b_2 + b_3 + b_4 2^{-1} + b_5 2^{-2} + \cdots + b_{16} 2^{-12})$, or

$$i_O = 2\frac{V_{REF}}{R_r}(b_1 2^{-1} + b_2 2^{-2} + \cdots + b_{16} 2^{-16}) \qquad (12.15)$$

indicating a 16-bit conversion with $I_{FSR} = 2$ mA. Two popular examples of 16-bit monolithic DACs utilizing this architecture are the PCM52/53 and HI-DAC16.

12.3 MULTIPLYING DAC APPLICATIONS

The R-$2R$ ladder DACs of Figs. 12.11 and 12.12 are especially suited to monolithic fabrication in CMOS technology.[6] The switches are implemented with CMOS transistors, and the ladder and the feedback resistor $R_f = R$ are fabricated by thin-film deposition on the CMOS die. Because of process variations, the resistances, though highly matched, are not necessarily accurate. For instance, a ladder with a nominal rating of 10 kΩ may in practice lie in the range of 5 kΩ to 20 kΩ.

Figure 12.21 shows the circuit diagram of the kth switch, $k = 1, 2, \ldots, n$. The switch proper consists of the n-MOS pair M_8-M_9, while the remaining FETs accept

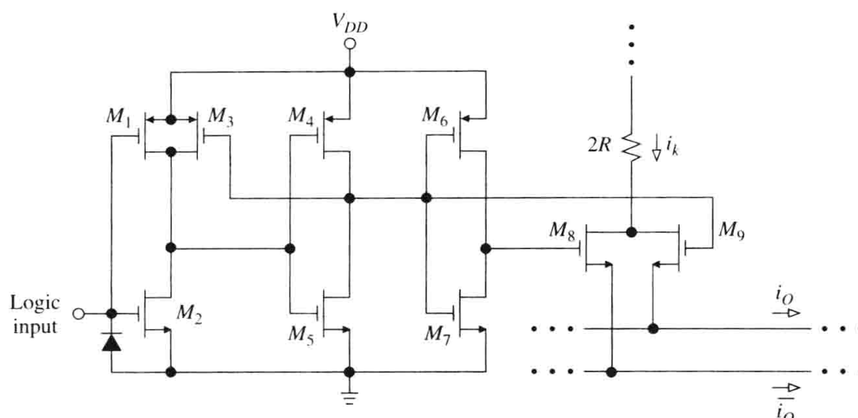


FIGURE 12.21
CMOS switch for R-$2R$ ladder.

TTL- and CMOS-compatible logic inputs to provide antiphase gate drives for M_8 and M_9. When the logic input is high, M_8 is off and M_9 is on, so i_k is diverted to the i_O bus. When the input is low, M_8 is on, M_9 is off, and i_k is now diverted to the $\overline{i_O}$ bus.

The nonzero resistance $r_{ds(\text{on})}$ of the switches tends to disrupt the 2:1 ratio of the ladder resistances and degrade performance. Since $r_{ds(\text{on})}$ is proportional to the ratio of the channel length L to the channel width W, it could be minimized by fabricating M_8 and M_9 with $L/W \ll 1$; this, however, would lead to extravagant device geometries. A common technique for overcoming this drawback is to taper switch geometries to achieve, at least in the MS bit positions, binary-weighted switch resistances such as $r_{ds1(\text{on})} = 20\ \Omega$, $r_{ds2(\text{on})} = 40\ \Omega$, $r_{ds3(\text{on})} = 80\ \Omega$, and so on. Since the currents halve as the switch resistances double, the product $r_{dsk(\text{on})} \times i_k$ remains constant throughout the tapered bit positions, causing a systematic switch voltage drop, whose value is typically 10 mV. Since this drop is effectively being subtracted from V_{REF}, the result is a gain error that is readily trimmed by adjusting R_f.

EXAMPLE 12.3. A CMOS DAC with $n = 12$ is operated in the current mode depicted in Fig. 12.11. If $V_{\text{REF}} = 10.0$ V and the DAC is calibrated at 25 °C, specify TC(V_{REF}) and TC(V_{OS}) so that the individual drift errors contributed by the reference and the op amp are less than $\pm\frac{1}{4}$ LSB over the operating range of 0 °C to 70 °C.

Solution. We have $\frac{1}{4}$ LSB $= 10.0/2^{14} = 0.61$ mV. Since the maximum temperature excursion from the point of calibration is $70° - 25° = 45$ °C, the individual drifts must not exceed $\pm 0.61 \times 10^{-3}/45 \cong \pm 13.6\ \mu$V/°C. This gives $\text{TC}_{\max}(V_{\text{REF}}) = \pm 1.36$ ppm/°C. Moreover, using a conservative estimate of 2 V/V for the noise gain of the op amp, we have $\text{TC}_{\max}(V_{OS}) \cong \pm 13.6/2 = \pm 6.8\ \mu$V/°C.

In the following we shall use the functional diagram of Fig. 12.22 to represent a CMOS DAC. This structure is available from various manufacturers in a range of resolutions (8 to 14 bits) and configurations (single, dual, quad, and octal packages). Many versions include input buffer latches to facilitate microprocessor interfacing.

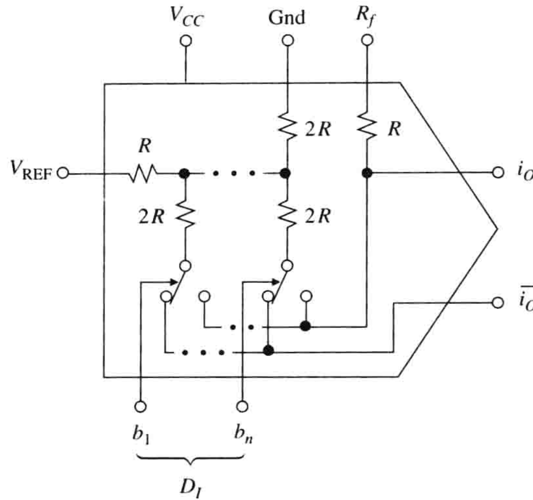


FIGURE 12.22
Functional diagram of a multiplying DAC.

Depending on resolution, settling times range from under 100 ns to over 1 μs. One of the earliest and most popular families of CMOS DACs is the AD7500 series.

MDAC Applications

The reference voltage of a CMOS DAC can be varied over positive as well as negative values, including zero. This inherent multiplicative ability makes CMOS DACs, aptly called *MDACs,* suited to a variety of digitally programmable applications.[6]

The circuits of Fig. 12.23 provide, respectively, digitally programmable attenuation and amplification. Using Eq. (12.10) with $R_f = R$, we find that the attenuator

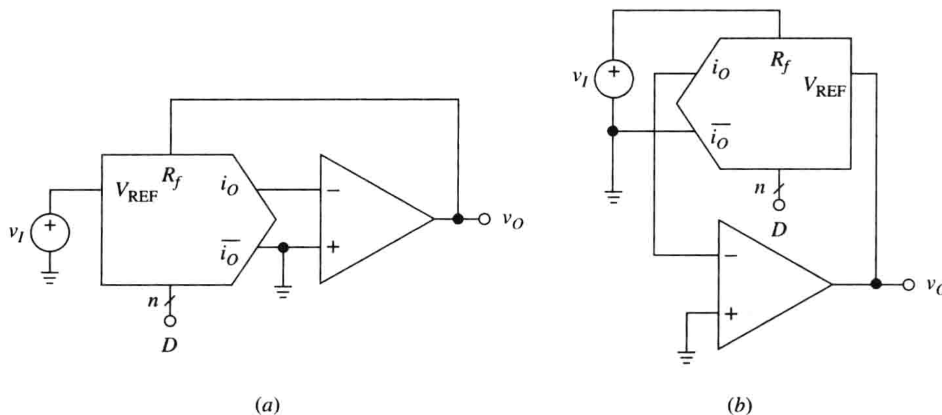


FIGURE 12.23
(*a*) Digitally programmable attenuator: $v_O = -Dv_I$; (*b*) digitally programmable amplifier: $v_O = (-1/D)v_I$.

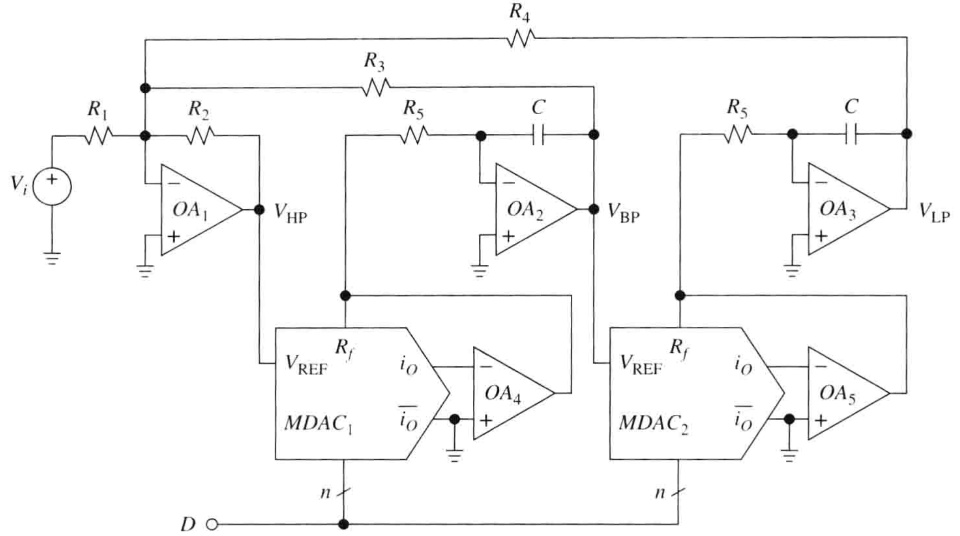


FIGURE 12.24
Digitally programmable filter.

of Fig. 12.23*a* gives $v_O = -Dv_I$, so its gain $A = -D$ is programmable from 0 to $-(1-2^{-n})$ V/V $\cong -1$ V/V in steps of 2^{-n} V/V. In the amplifier of Fig. 12.23*b* we have $v_I = -Dv_O$, or $v_O = (-1/D)v_I$. Its gain $A = -1/D$ is programmable from $-1/(1-2^{-n}) \cong -1$ V/V when all bits are 1, to -2 V/V when $b_1 b_2 \dots b_n = 10\dots0$, to 2^n V/V when $b_1 \dots b_{n-1} b_n = 0\dots01$, to the full open-loop gain a when all bits are 0. To combat the effect of the stray capacitance of the i_O bus, it is advisable to connect a stabilizing capacitance C_f of a few tens of picofarads between the output and the inverting input of the op amp.[5]

If we cascade the attenuator of Fig. 12.23*a* with a Miller integrator having unity-gain frequency ω_1, the transfer function of the composite circuit is $H = (-D) \times [-1/(j\omega/\omega_1)] = 1/(j\omega/D\omega_1)$. This represents a noninverting integrator with a digitally programmable unity-gain frequency of $D\omega_1$. Such an integrator can be used to implement a digitally programmable filter. The filter example of Fig. 12.24 is a state-variable topology of the type encountered in Fig. 4.37, so we can reuse Eq. (4.34) and write

$$\omega_0 = D\sqrt{R_2/R_4}/R_5C \qquad Q = R_3/\sqrt{R_2R_4} \tag{12.16a}$$

$$H_{0\mathrm{HP}} = -R_2/R_1 \qquad H_{0\mathrm{BP}} = -R_3/R_1 \qquad H_{0\mathrm{LP}} = -R_4/R_1 \tag{12.16b}$$

indicating that we can program ω_0 digitally from $2^{-n}\sqrt{R_2/R_4}/R_5C$ to $(1-2^{-n})\sqrt{R_2/R_4}/R_5C$. Once we have a digitally programmable filter, we can readily turn it into a digitally programmable oscillator by letting $Q \to \infty$ (see Problem 12.12).

EXAMPLE 12.4. In the circuit of Fig. 12.24 specify suitable components for $Q = 1/\sqrt{2}$, $H_{0\mathrm{BP}} = -1$ V/V, and f_0 digitally programmable in 10-Hz steps by means of 10-bit MDACs.

Solution. Impose $R_2 = R_4 = 10.0$ kΩ, and let $C = 1.0$ nF. Then, the full-scale range is $f_{0(FSR)} = 2^{10} \times 10 = 10.24$ kHz, so $R_5 = 1/(2\pi 10,240 \times 10^{-9}) = 15.54$ kΩ (use 15.4 kΩ, 1%).

Use fast op amps with low-input-offset error and noise characteristics and wide dynamics, such as the OPA627 JFET-input op amps. To avoid high-frequency Q enhancement, phase-error compensation may be required, as discussed in Section 6.5.

Figure 12.25 shows a digitally programmable waveform generator. The circuit is similar to that encountered in Fig. 10.19*a*, except for the use of an MDAC to control the rate of capacitance charge/discharge digitally. To avoid the uncertainties of the ladder resistances, the MDAC is current-driven using the REF200 100-μA current source. When v_{SQ} is high, I_{REF} enters the MDAC; when v_{SQ} is low, I_{REF} exits the MDAC. In either case the MDAC divides this current to give $i_O = \pm DI_{REF}$. To find the frequency of oscillation f_0, apply Eq. (10.2) with $\Delta t = 1/(2f_0)$, $I = DI_{REF}$, and $\Delta v = 2V_T = 2(R_1/R_2)V_{clamp}$, where $V_{clamp} = 2V_{D(on)} + V_{Z5}$. The result is

$$f_0 = D\frac{(R_2/R_1)I_{REF}}{4CV_{clamp}} \tag{12.17}$$

indicating that f_0 is linearly proportional to D.

EXAMPLE 12.5. In the circuit of Fig. 12.25 specify suitable components for 5-V waveform amplitudes and f_0 digitally programmable in 1-Hz steps by means of a 12-bit MDAC.

Solution. For $V_{clamp} = 5$ V, use $V_{Z5} = 3.6$ V. Moreover, use $R_1 = R_2 = 20$ kΩ and $R_3 = 6.2$ kΩ. The full-scale range is $f_{0(FSR)} = 2^{12} \times 1 = 4.096$ kHz, so Eq. (12.17) gives $C = 100 \times 10^{-6}/(20 \times 4096) = 1.22$ nF (use 1.0 nF, which is more easily available, and raise R_1 to 24.3 kΩ, 1%). Use a low-offset JFET-input op amp for *OA*, and a high slew-rate op amp for *CMP*.

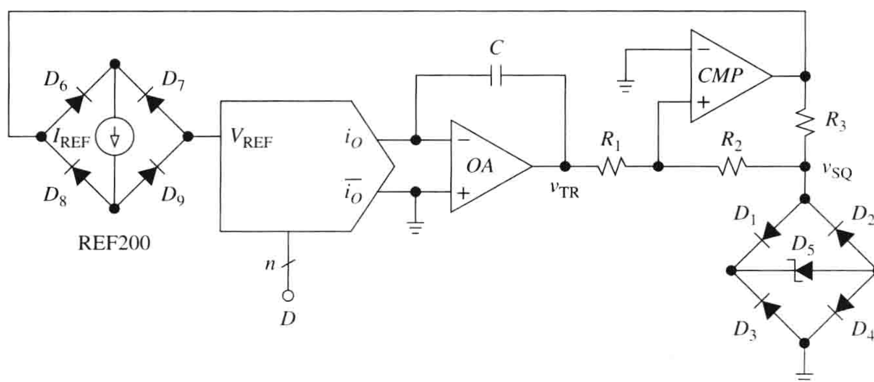


FIGURE 12.25
Digitally programmable triangular/square-wave oscillator.

12.4
A-D CONVERSION TECHNIQUES

This section discusses popular ADC techniques, such as DAC-based ADCs, flash ADCs, integrating ADCs, and variants thereof.[3,4] A more recent technique, known as sigma-delta (Σ-Δ) conversion, is addressed in Section 12.5.

DAC-Based A-D Conversion

A-D conversion can be accomplished by using a DAC and a suitable register to adjust the DAC's input code until the DAC's output comes within $\pm\frac{1}{2}$ LSB of the analog input. The code that achieves this is the desired ADC output $b_1 \ldots b_n$. As shown in Fig. 12.26, this technique requires suitable logic circuitry to direct the register to perform the code search on the arrival of the START command, and a voltage comparator to announce when v_O has come within $\pm\frac{1}{2}$ LSB of v_I and thus issue an end-of-conversion (EOC) command. Moreover, to center the analog range properly, the DAC output must be offset by $+\frac{1}{2}$ LSB, per Fig. 12.5*b*.

The simplest code search is a *sequential search,* obtained by operating the register as a binary counter. As the counter steps through consecutive codes starting from $0 \ldots 0$, the DAC produces an increasing staircase, which the comparator then compares against v_I. As soon as this staircase reaches v_I, *CMP* fires and stops the counter. This also serves as an EOC command to notify that the desired code is sitting in the counter. The counter must be stepped at a low enough frequency to allow for the DAC to settle within each clock cycle. Considering that a conversion can take as many as $2^n - 1$ clock periods, this technique is limited to low-speed applications. For example, a 12-bit ADC with a 1-MHz counter clock will take $(2^{12} - 1)\ \mu\text{s} = 4.095$ ms to convert a full-scale input.

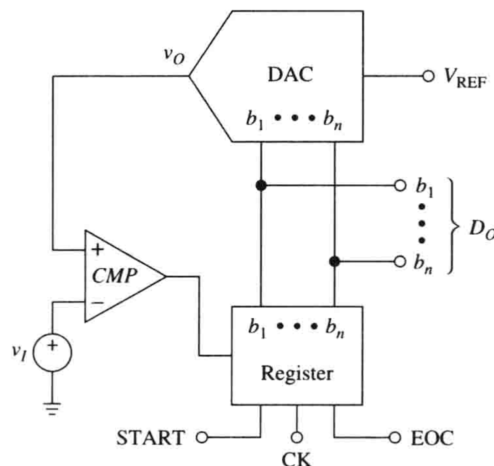


FIGURE 12.26
Functional diagram of a DAC-based ADC.

A better approach is to allow the counter to start counting from the most recent code rather than restarting from zero. If v_I has not changed drastically since the last conversion, fewer counts will be needed for v_O to catch up with v_I. Also referred to as a *tracking* or a *servo converter*, this scheme uses the register as an up/down counter with the count direction controlled by the comparator: counting will be up when $v_O < v_I$, and down when $v_O > v_I$. Whenever v_O crosses v_I, the comparator changes state and this is taken as an EOC command. Clearly, conversions will be relatively fast only as long as v_I does not change too rapidly between consecutive conversions. For a full-scale change, the conversion will still take $2^n - 1$ clock periods.

The fastest code-search strategy uses binary search techniques to complete an n-bit conversion in just n clock periods, regardless of v_I. Following is a description of two implementations: the *successive-approximation* and the *charge-redistribution* ADCs.

Successive-Approximation Converters (SA ADCs)

This technique uses the register as a *successive-approximation register* (SAR) to find each bit by trial and error. Starting from the MSB, the SAR inserts a trial 1 and then interrogates the comparator to find whether this causes v_O to rise above v_I. If it does, the trial bit is changed back to 0; otherwise it is left as 1. The procedure is then repeated for all subsequent bits, one bit at a time, in a way similar to a chemist's balance. Figure 12.27 illustrates how a 10.8-V input is converted to a 4-bit code with $V_{FSR} = 16$ V. The analog range, in volts, is at the left, and the digital codes at the right. To ensure correct results, the DAC output must be offset by $-\frac{1}{2}$ LSB, or -0.5 V in our example. The conversion takes place as follows.

Following the arrival of the START command, the SAR sets b_1 to 1 with all remaining bits at 0 so that the trial code is 1000. This causes the DAC to output

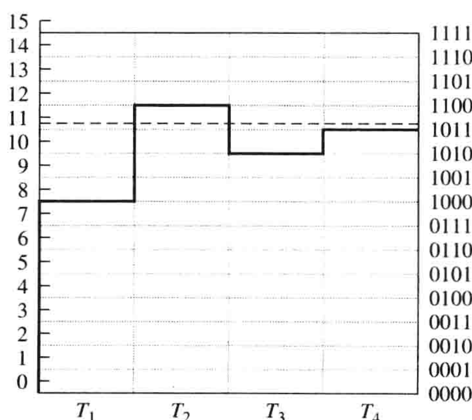


FIGURE 12.27
Idealized DAC output for the 4-bit successive-approximation conversion of $v_I = 10.8$ V with $V_{FSR} = 16$ V.

$v_O = 16(1 \times 2^{-1} + 0 \times 2^{-2} + 0 \times 2^{-3} + 0 \times 2^{-4}) - 0.5 = 7.5$ V. At the end of clock period T_1, v_O is compared against v_I, and since $7.5 < 10.8$, b_1 is left at 1.

At the beginning of T_2, b_2 is set to 1, so the trial code is now 1100 and $v_O = 16(2^{-1} + 2^{-2}) - 0.5 = 11.5$ V. Since $11.5 > 10.8$, b_2 is changed back to 0 at the end of T_2.

At the beginning of T_3, b_3 is set to 1, so the trial code is 1010 and $v_O = 10 - 0.5 = 9.5$ V. Since $9.5 < 10.8$, b_3 is left at 1.

At the beginning of T_4, b_4 is set to 1, so the trial code is 1011 and $v_O = 11 - 0.5 = 10.5$ V. Since $10.5 < 10.8$, b_4 is left at 1. Thus, when leaving T_4, the SAR has generated the code 1011, which ideally corresponds to 11 V. Note that any voltage in the range $10.5 \text{ V} < v_I < 11.5$ V would have led to the same code.

Since the entire conversion takes a total of n clock cycles, an SA ADC offers a major speed improvement over a sequential-search ADC. For instance, a 12-bit SA ADC with a clock frequency of 1 MHz will complete a conversion in 12 μs.

Figure 12.28 shows an actual implementation[7] using the Am2504 SAR and the Am6012 bipolar DAC, whose settling time is 250 ns, along with the CMP-05 comparator, whose response time to a 1.2-mV overdrive ($\frac{1}{2}$ LSB) is 125 ns maximum.

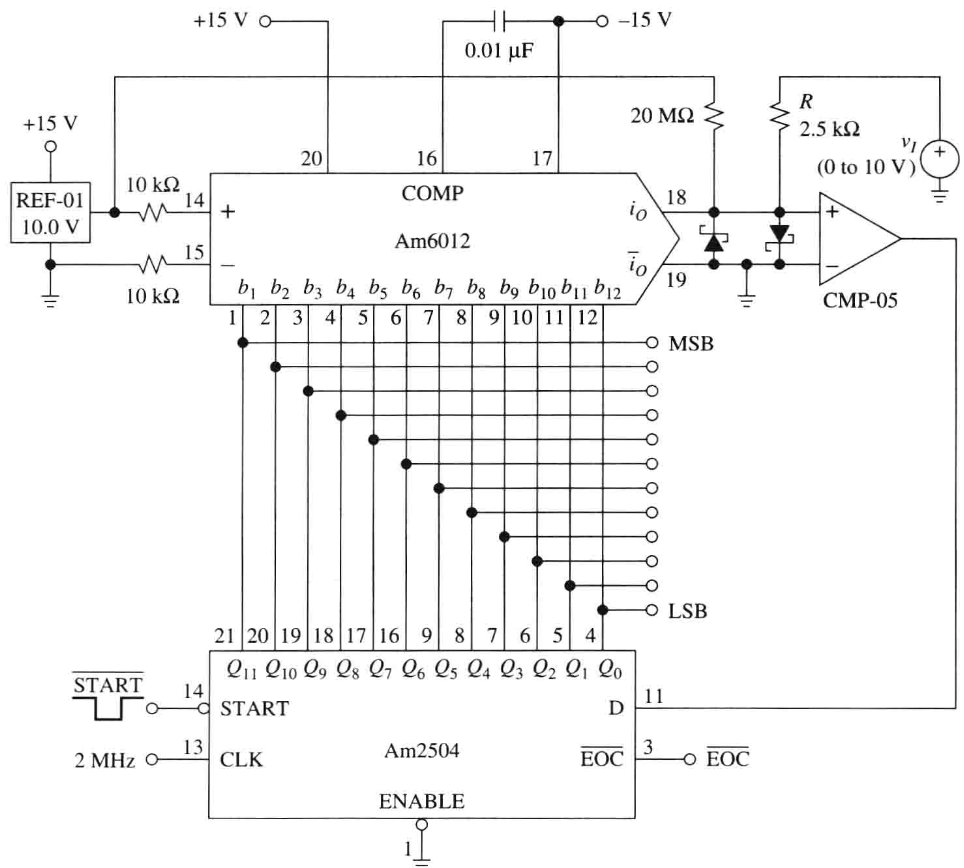


FIGURE 12.28
12-bit, 6-μs successive-approximation ADC.

The desired output code is available both in parallel form from Q_0 through Q_{11}, or in serial form at the data pin D.

To take full advantage of the bipolar DAC speed, i_O is converted to a voltage for the comparator via simple resistive termination. Since its input is $v_D = v_I - Ri_O$, the comparator is in effect comparing i_O against v_I/R. The function of the 20-MΩ resistance is to provide the required $-\frac{1}{2}$-LSB shift, and that of the Schottky diodes is to limit the voltage swing at the comparator input in order to reduce delays caused by the stray output capacitance of the DAC.

The primary factors affecting the speed of an SA ADC are the settling time of the DAC and the response time of the comparator. The conversion time can be further reduced by a number of ingenious techniques,[7] such as comparator speed-up techniques, or variable-clock techniques, which exploit the faster settling times in the least significant bit positions.

The resolution of an SA ADC is limited by the resolution and linearity of the DAC, and the gain of the comparator. A crucial requirement is that the DAC be monotonic to prevent the occurrence of missing codes. The comparator, besides adequate speed, must provide enough gain to magnify an LSB step to a full output logic swing, or $a \geq (V_{OH} - V_{OL})/(V_{FSR}/2^n)$. For instance, with $V_{OH} = 5$ V, $V_{OL} = 0$ V, $V_{FSR} = 10$ V, and $n = 12$, we need $a \geq 2048$ V/V. Another important requirement is that during conversion v_I remain constant within $\pm\frac{1}{2}$ LSB; otherwise an erroneous code may result. For instance, if v_I were to rise above 11.5 V after the second clock period in Fig. 12.27, there would be no way for the SAR to go back and change b_2, so a wrong output code would result. This is avoided by preceding the ADC with a suitable SHA.

SA ADCs are available from a variety of sources and in a wide range of performance characteristics and prices. Conversion times typically range from under 1 μs for the faster 8-bit units to tens of microseconds for the high-resolution ($n \geq 14$) types. SA ADCs equipped with an on-chip SHA are referred to as *sampling* ADCs. A popular example is the AD1674 12-bit, 100-kilosamples per second (ksps) SA ADC.

Charge-Redistribution Converters (CR ADCs)

The circuit of Fig. 12.29 performs a successive-approximation conversion using a weighted-capacitor DAC of the type of Fig. 12.8. Its operation involves three cycles called the *sample, hold,* and *redistribution* cycles.

During the sample cycle, SW_0 grounds the top-plate bus while SW_i and SW_1 through SW_{n+1} connect the bottom plates to v_I, thus precharging the entire capacitor array to v_I.

During the hold cycle, SW_0 is opened and the bottom plates are switched to ground, thus causing the top-plate voltage to swing to $-v_I$. The voltage presented to the comparator at the end of this cycle is thus $v_P = -v_I$.

During the redistribution cycle, SW_0 is still open, SW_i is connected to V_{REF}, and the remaining switches are sequentially flipped from ground to V_{REF}, and possibly back to ground, to perform a successive-approximation search for the desired code.

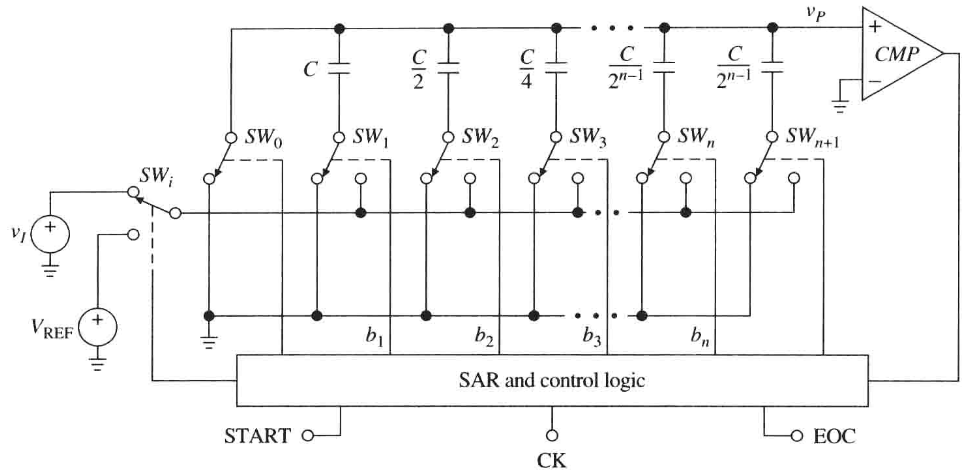


FIGURE 12.29
Charge-redistribution ADC.

Flipping a given switch SW_k from ground to V_{REF} causes v_P to increase by the amount $V_{REF}(C/2^{k-1})/C_t = V_{REF}2^{-k}$. If it is found that this increase causes the comparator to change state, then SW_k is returned to ground; otherwise it is left at V_{REF} and the next switch is tried. This procedure is repeated at each bit position, starting from the MSB and progressing down to the LSB (excluding the terminating capacitor switch, which is left permanently grounded). It is readily seen that at the end of the search the voltage presented to the comparator is

$$v_P = -v_I + V_{REF}(b_1 2^{-1} + b_2 2^{-2} + \cdots + b_n 2^{-n})$$

and that v_P is within $\pm\frac{1}{2}$ LSB of 0 V. Thus, the final switch pattern provides the desired output code.

Because of the exponential increase of capacitance spread with n, practical CR ADCs are limited to $n \leq 10$. One way to increase resolution is to combine charge redistribution with potentiometric techniques,[2] as exemplified in Fig. 12.30. Here a resistor string partitions V_{REF} into 2^{n_H} inherently monotonic voltage segments, and an n_L-bit weighted-capacitor DAC interpolates within the selected segment. As long as the capacitances are ratio-accurate to n_L bits, the composite DAC will retain monotonicity to $n = n_H + n_L$ bits, so using it as part of an SA conversion will avoid missing codes. A conversion proceeds as follows.

Initially, SW_f is closed to autozero the comparator, and the bottom plates are connected via the L bus and SW_L to the analog input v_I. This precharges the capacitor array to v_I minus the comparator's threshold voltage, thus removing this threshold as a possible source of error.

Next, SW_f is opened, and an SA search among the resistor string taps is performed to find the segment within which the voltage held in the capacitor array lies. The outcome of this search is the n_H-bit portion of the desired code.

Once the segment has been found, the H and L busses are connected to the extremes of the corresponding resistor, and a second SA search is performed to find the individual bottom-plate switch settings that make the comparator input converge

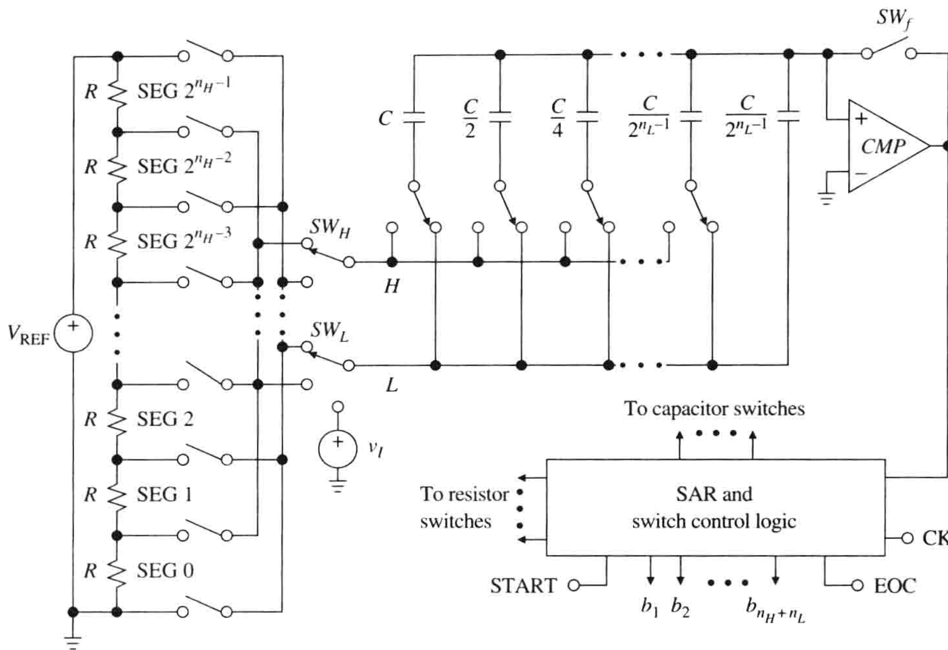


FIGURE 12.30
High-resolution charge-redistribution ADC.

to its threshold. The outcome of this search is the n_L-bit portion of the desired code. For instance, with $n_H = 4$ and $n_L = 8$, the circuit provides 12 bits of resolution without excessive demands in terms of circuit complexity or capacitance spread and matching.

Before concluding, we wish to point out that the analog input, conveniently represented in terms of an ideal voltage source v_I in both Figs. 11.29 and 11.30, is likely to be the output of a buffer/amplifier, whose nonzero output impedance may cause conversion inaccuracies, not to mention the stability problems associated with capacitive loads. A common cure is to interpose a suitable decoupling *R*-*C* network between the driver and the ADC (consult the literature and application notes for guidelines on the optimal selection of the *R*-*C* values[2]).

Flash Converters

The circuit of Fig. 12.31 uses a resistor string to create $2^n - 1$ reference levels separated from each other by 1 LSB, and a bank of $2^n - 1$ high-speed latched comparators to simultaneously compare v_I against each level. Note that to position the analog signal range properly, the top and bottom resistors must be $1.5R$ and $0.5R$, as shown. As the comparators are strobed by the clock, the ones whose reference levels are below v_I will output a logic 1, and the remaining ones a logic 0. The result, referred to as a *bar graph,* or also as a *thermometer* code, is then converted to the desired output code $b_1 \ldots b_n$ by a suitable decoder, such as a priority encoder. Since input sampling and latching take place during the first phase of the clock

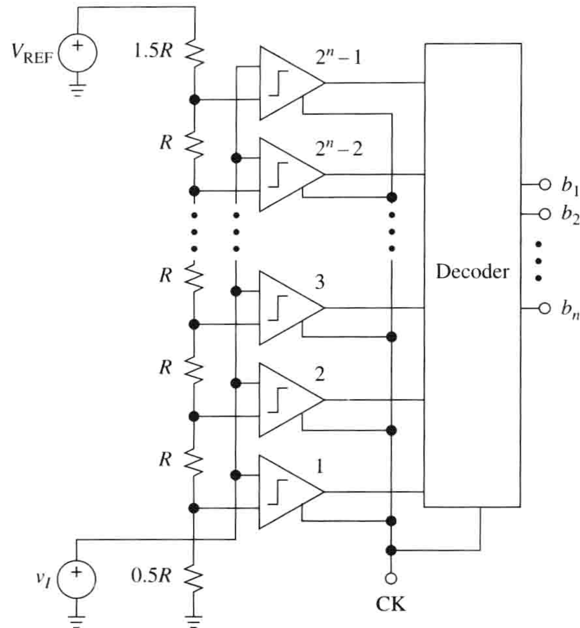


FIGURE 12.31
n-bit flash converter.

period, and decoding during the second phase, the entire conversion takes only one clock cycle, so this ADC is the fastest possible. Aptly called a *flash converter,* it is used in high-speed applications, such as video and radar signal processing, where conversion rates on the order of millions of samples per second (Msps) are required, and SA ADCs are generally not fast enough.

The high-speed and inherent-sampling advantages of flash ADCs are offset by the fact that $2^n - 1$ comparators are required. For instance, an 8-bit converter requires 255 comparators. The exponential increase with n in die area, power dissipation, and stray input capacitance makes flash converters impractical for $n > 10$. Flash ADCs are available in bipolar or in CMOS technology, with resolutions of 6, 8, and 10 bits, sampling rates of tens to hundreds of Msps, depending on resolution, and power dissipation ratings on the order of 1 W or less. Consult the catalogs to familiarize yourself with the range of available products.

Subranging Converters

Subranging ADCs trade speed for circuit complexity by splitting the conversion into two subtasks, each requiring less complex circuitry. Also called a *two-step,* or a *half-flash converter,* this architecture uses a coarse flash ADC to provide an n-bit accurate digitization of the n_H most-significant bits. These bits are then fed to a high-speed, n-bit accurate DAC to provide a coarse approximation to the analog input. The difference between this input and the DAC output, called the *residue,* is magnified by 2^{n_H} V/V by an amplifier called the *residue amplifier* (RA), and finally fed to a fine flash ADC for the digitization of the n_L least-significant bits of the n-bit

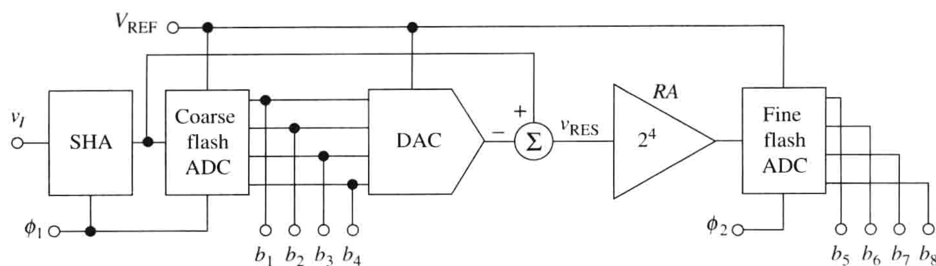


FIGURE 12.32
8-bit subranging ADC. (Note that DAC must be 8-bit accurate.)

code, where $n = n_H + n_L$. Note that the half-flash requires an SHA to hold the value of v_I during the digitization of the residue.

Figure 12.32 exemplifies an 8-bit converter with $n_L = n_H = 4$. Besides the SHA, the DAC, and the RA, the circuit uses $2(2^4 - 1) = 30$ comparators, indicating a substantial saving compared to the 255 comparators required by a full-flash. (This saving is even more dramatic for $n \geq 10$.) The main price for this saving is a longer conversion time, with the first phase comprising the conversion time of the coarse ADC, the acquisition time of the SHA, and the settling time of the DAC-subtractor-RA block, and the second phase comprising the conversion time of the fine ADC. Moreover, the requirement that the DAC be n-bit accurate may be a heavy requirement.

Subranging ADCs, though not as fast as full-flash ADCs, are still comparably faster than SA ADCs, so the subranging architecture, or variants[4] thereof, is used in a number of high-speed ADC products.

Pipelined Converters

Pipelined ADCs break down the conversion task into a sequence of N serial subtasks, and use SHA interstage isolation to allow for the individual subtasks to proceed concurrently to achieve high throughput rates. With reference to Fig. 12.33, each subtask stage consists of an SHA, an ADC, a DAC, a subtractor, and an RA, with

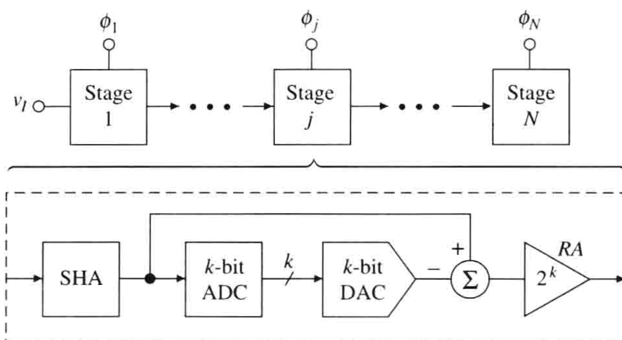


FIGURE 12.33
Pipeline ADC architecture.

some or even all functions often combined in one circuit.[4] The first stage samples v_I, digitizes k bits, and uses a DAC-subtractor-RA circuit to create a residue for the next stage in the pipeline. The next stage samples the incoming residue and performs a similar sequence of operations while the previous stage begins processing the next sample. The ability of the various stages to operate concurrently makes the conversion rate depend on the speed of only one stage, usually the first stage. Pipelined structures are used in a variety of formats, including the case $k = 1$, which results in the simplest per-stage circuitry, though n such stages are needed. However, if stages are reused, considerable savings in die area can be achieved.

Integrating-Type Converters

These converters perform A-D conversion indirectly by converting the analog input to a linear function of time and thence to a digital code. The two most common converter types are the *charge-balancing* and *dual-slope* ADCs.

Charge-balancing ADCs convert the input signal to a frequency, which is then measured by a counter and converted to an output code proportional to the analog input.[8] These converters are suited to applications where it is desired to exploit the ease with which a frequency is transmitted in noisy environments or in isolated form, such as telemetry. However, as seen in Section 10.7, the transfer characteristic of a VFC depends on an RC product whose value is not easily maintained with temperature and time. This drawback is ingeniously overcome by dual-slope converters.

As shown in the functional diagram of Fig. 12.34, a dual-slope ADC, also called a *dual-ramp* ADC, is based on a high-input-impedance buffer, a precision integrator, and a voltage comparator. The circuit first integrates the input signal v_I for a fixed duration of 2^n clock periods, and then it integrates an internal reference V_{REF} of opposite polarity until the integrator output is brought back to zero. The number N

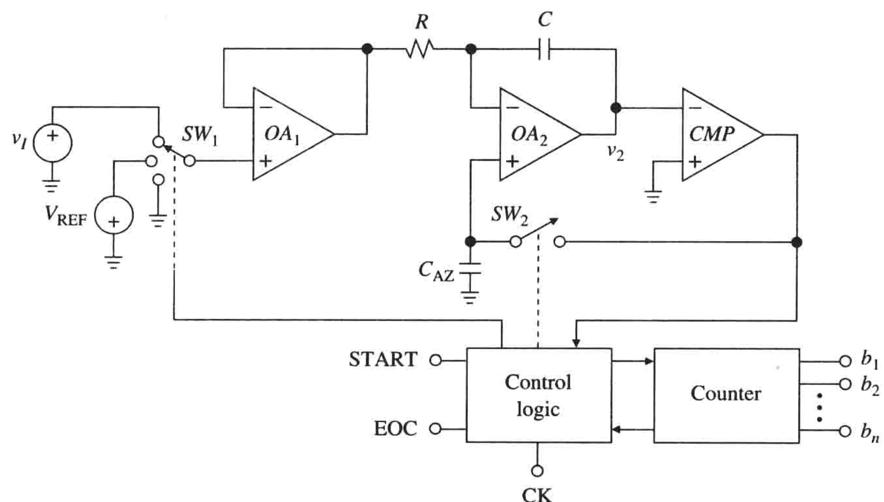


FIGURE 12.34
Functional diagram of a dual-slope ADC.

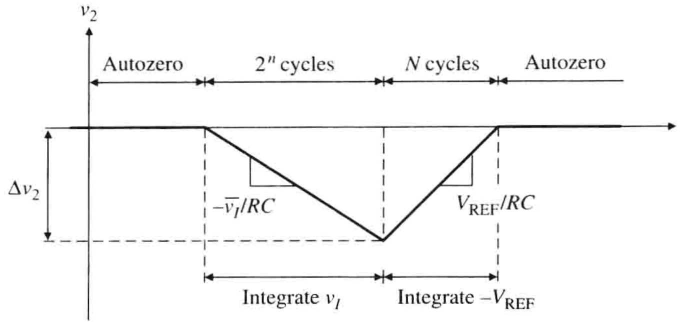


FIGURE 12.35
Dual-slope waveform.

of clock cycles required to return to zero is proportional to the value of v_I averaged over the integration period. Consequently, N represents the desired output code. With reference to the waveform diagram of Fig. 12.35, following is a detailed description of how the circuit operates.

Prior to the arrival of the START command, SW_1 is connected to ground and SW_2 closes a loop around the integrator-comparator combination. This forces the autozero capacitance C_{AZ} to develop whatever voltage is needed to bring the output of OA_2 right to the comparator's threshold voltage and leave it there. This phase, referred to as the *autozero phase,* provides simultaneous compensation for the input offset voltages of all three amplifiers. During the subsequent phases, when SW_2 opens, C_{AZ} acts as an analog memory to hold the voltage required to keep the net offset nulled.

At the arrival of the START command, the control logic opens SW_2, connects SW_1 to v_I (which we assume to be positive), and enables the counter, starting from zero. This phase is called the *signal integrate phase*. As the integrator ramps downward, the counter counts until, 2^n clock periods later, it overflows. This marks the end of the current phase. The swing Δv_2 described by the integrator during this interval is found via Eq. (10.2) as $C\Delta v_2 = (\overline{v_I}/R) \times 2^n \times T_{CK}$, where T_{CK} is the clock period, and $\overline{v_I}$ the average of v_I over $2^n T_{CK}$.

As the overflow condition is reached, the counter resets automatically to zero and SW_1 is connected to $-V_{REF}$, causing v_2 to ramp upward. This is called the *deintegrate phase.* Once v_2 again reaches the comparator threshold, the comparator fires to stop the counter and issues an EOC command. The accumulated count N is such that $C\Delta v_2 = (V_{REF}/R)NT_{CK}$. Since $C\Delta v_2$ is the same during the two phases, we get

$$N = 2^n \frac{\overline{v_I}}{V_{REF}} \tag{12.18}$$

We make a number of important observations.

1. The conversion accuracy is independent of R, C, T_{CK}, and the input offset voltage of the three amplifiers. As long as these parameters remain stable over the conversion period, they affect the two integration phases equally, so long-term drifts are automatically eliminated.

2. An integrating ADC offers excellent linearity and resolution, and virtually zero differential nonlinearity. With an integrator of suitable quality, nonlinearity errors can be kept below 0.01%, and resolution can be pushed above 20 bits. Moreover, since v_2 is a continuous function of time, differential nonlinearity, within the limits of clock jitter, is virtually absent, so there are no missing codes.
3. A dual-slope ADC provides excellent rejection of ac noise components with frequencies that are integral multiples of $1/(2^n T_{CK})$. For instance, if we specify T_{CK} so that $2^n T_{CK}$ is a multiple of $1/60 = 16.67$ ms, then any 60-Hz pickup noise superimposed on the input signal will be averaged to zero. In particular, if $2^n T_{CK} = 100$ ms, the ADC will reject both 50-Hz and 60-Hz noise.
4. An integrating converter does not require an SHA at the input. If v_I changes, the converter will simply average it out over the signal-integrate period.

The main drawback of dual-slope ADCs is a low conversion rate. For instance, imposing $2^n T_{CK} = 1/60$ and allowing as many clock periods to complete the deintegrate phase for a full-scale input, it follows that the conversion rate is less than 30 sps. These converters are suited to highly accurate measurements of slowly varying signals, as in thermocouple measurements, weighing scales, and digital multimeters.

Dual-slope ADC ICs are available from a variety of sources, usually in CMOS technology. Besides autozero capabilities, they offer automatic input polarity sensing and reference polarity switching to provide sign and magnitude information. Moreover, they are available both in microprocessor-compatible and in display-oriented versions. The latter provides the output code in a format suitable for driving decimal LCD or LED displays, and their resolution is expressed in terms of decimal digits rather than bits. Since the leftmost digit is usually allowed to run only to unity, it is counted as $\frac{1}{2}$ digit. Thus, a $4\frac{1}{2}$-digit sign-plus-magnitude ADC having $V_{FSR} = 200$ mV yields all decimal codes within the range of ± 199.99 mV and with a resolution of 10 μV. An example is the ICL7129 $4\frac{1}{2}$-digit ADC, which, with the help of suitable support circuitry, is easily turned into a full-fledged multimeter to measure both dc and ac voltages and currents, as well as resistances.

We are now able to compare the circuit complexity and the required clock-cycles for the architectures discussed so far:

	Flash	Pipeline	SA	Integrating
Complexity :	2^n	n	1	1
Conversion :	1	1	n	2^n

12.5 OVERSAMPLING CONVERTERS

It is apparent that the most critical part of a data converter is its analog circuitry. Because of component mismatches and nonlinearities, drift and aging, noise, dynamic limitations and parasitics, resolution and speed can be pushed only so far. Oversampling converters ease analog-circuitry requirements at the expense of more complex digital circuitry. These converters are ideal for mixed-mode IC fabrication processes, where fast digital-processing circuitry is far more easily implemented than precise analog circuitry. The principal benefits of oversampling followed by

digital filtering are *relaxed analog-filter requirements* and *quantization-noise reduction.* Sigma-delta (Σ-Δ) converters combine with these benefits the additional benefit of *noise shaping* to achieve truly high resolutions ($\geq$ 16 bits) with the simplest analog circuitry (1-bit digitizers).

Before embarking on the study of oversampling and noise shaping, we need to examine in greater detail conventional sampling, also referred to as *Nyquist-rate sampling.*

Nyquist-Rate Sampling

The digitization process, depicted in Fig. 12.36*a*, has a profound impact on the frequency spectrum of the input signal. We are primarily interested in the situation from dc to the sampling frequency f_S. As depicted in Fig. 12.36*b*, this range consists of two zones, namely, zone I extending from dc to $f_S/2$, and zone II extending from $f_S/2$ to f_S. Zone I is also called the *baseband,* and $f_S/2$ is called the *Nyquist bandwidth.* The effects of digitization are twofold:[1]

1. Digitization, viewed as *discretization in time,* creates additonal spectral components, called *images,* at locations symmetric about the midpoint $f_S/2$; for instance, a spectral component of v_I at $f = f_I$ results in an image at $f = f_S - f_I$, as shown in Fig. 12.36*b*, top.
2. Digitization, viewed as *discretization in amplitude,* introduces quantization noise, as discussed in Section 12.1. The noise power of v_I folds into the baseband, in the manner depicted in Fig. 12.36*b*, bottom.

If v_I is a relatively active or busy signal, its quantization noise can, under certain conditions,[9,10] be treated as white noise with spectral density

$$e_q = \frac{q}{\sqrt{f_S/2}} \tag{12.19}$$

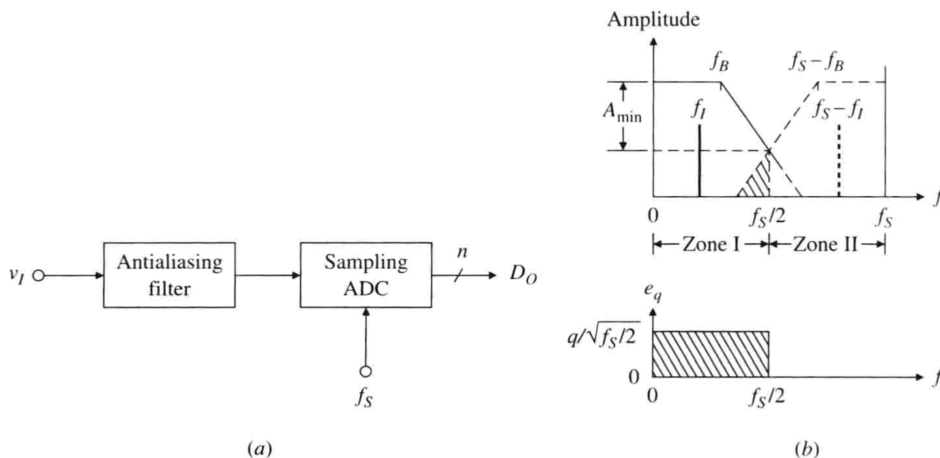


FIGURE 12.36
Nyquist sampling with analog filtering.

where $q = V_{FSR}/2^n\sqrt{12}$. The rms value is

$$E_q = \left(\int_0^{f_S/2} e_q^2\, df\right)^{1/2} = q \qquad (12.20)$$

or $E_q = V_{FSR}/2^n\sqrt{12}$, in accordance with Eq. (12.4). As we know, this results in

$$\text{SNR}_{\text{max}} = 6.02n + 1.76\text{ dB} \qquad (12.21)$$

With reference to Fig. 12.36*b*, top, we observe that as long as all spectral components of v_I lie within zone I, their images will be confined within zone II. Processing the spectrum of the digitized signal with a low-pass filter having a cutoff frequency of $f_S/2$ will pass the baseband components and block their images, thus allowing for the full recovery of the spectrum of v_I. This spectrum can, in turn, be used to reconstruct v_I itself. However, should v_I possess spectral components in zone II, their images will creep into zone I, overlapping the legitimate components there and causing nonlinear distortion. This phenomenon, referred to as *aliasing,* introduces an ambiguity that prevents the recovery of the spectrum of v_I. *Nyquist's criterion* states that if we want to recover or reconstruct a signal of a given bandwidth f_B from its digitized version, the sampling rate f_S must be such that

$$f_S > 2f_B \qquad (12.22)$$

where $2f_B$ is called the *Nyquist rate.* This requirement can be met either by band-limiting v_I below $f_S/2$, or by raising f_S above the Nyquist rate.

A familiar aliasing example is offered by the spoked wheels of a stagecoach in a 16-mm, 24-frames-per-second Western. As long as the coach travels slowly enough relative to the camera's sampling rate of 24 frames per second, its wheels will appear to be turning correctly. However, as the coach speeds up, a point is reached where the wheels will appear to be slowing down, indicating an alias, or unwanted frequency, near the upper end of the baseband. Speeding up further will lower the alias frequency until it reaches dc, where the wheels will appear to be still. Any speed increase beyond this point will result in a negative alias frequency, making the wheels appear to be turning backward! These aliasing effects could be avoided either by limiting the filming only to slow scenes, or by increasing the number of frames per second.

In practical ADCs, to avoid wasting digital data rate, f_S is usually specified not far above the Nyquist rate of $2f_B$. For example, digital telephony, where the band of interest is $f_B = 3.2$ kHz and thus $2f_B = 6.4$ kHz, uses $f_S = 8$ kHz. Likewise, compact-disc audio, where $f_B = 20$ kHz and $2f_B = 40$ kHz, uses $f_S = 44.1$ kHz. Even though f_S is not strictly equal to $2f_B$, these converters are loosely referred to as *Nyquist-rate* converters.

It is apparent that in order to prevent any noise or spurious input spectral components above $f_S/2$ from folding into the baseband, an antialiasing filter is required. Such a filter must provide a flat response up to f_B and must roll off rapidly enough thereafter to provide the desired amount of suppression at $f_S/2$ and beyond. The shaded area of Fig. 12.36*b*, top, represents the baseband aliases of the unsuppressed signal and noise components above $f_S/2$. The contribution from these aliases must be kept below $\frac{1}{2}$ LSB by suitable choice of A_{min}. Such a choice, in turn, depends on the noise distribution and the spectral makeup of v_I for $f \geq f_S/2$. It is apparent

that the performance requirements of the antialiasing filter can be quite stringent. Elliptic filters are a common choice for this task because of their sharp cutoff rate, if at the price of a nonlinear phase response.

Oversampling

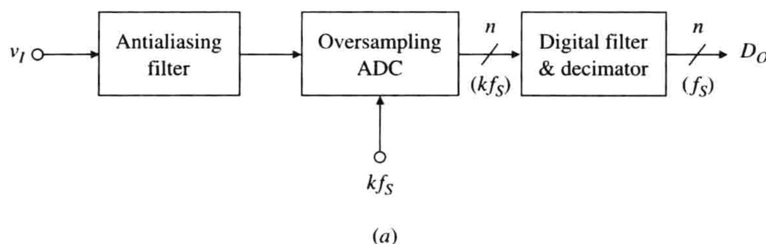
Consider now the effect of speeding up the sampling rate by a factor of k, $k \gg 1$. This is shown in Fig. 12.37*a*. The ensuing benefits, illustrated in Fig. 12.37*b*, are twofold:

1. The transition band of the analog filter preceding the digitizer is now much wider, providing an opportunity for a drastic reduction in circuit complexity. In fact, in oversampling converters of the Σ-Δ type, this filter can be as simple as a mere *RC* stage!
2. The quantization noise is now spread over a wider band, or

$$e_q = \frac{q}{\sqrt{kf_S/2}} \tag{12.23}$$

   indicating a spectral-density reduction by $\sqrt{k}$.

The price for the preceding benefits is the need for a *digital filter* at the output of the digitizer to (*a*) suppress any spectral components and noise above $f_S/2$, and



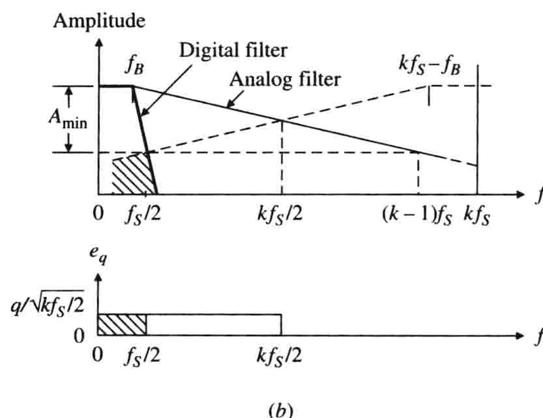


FIGURE 12.37
Oversampling with analog and digital filtering.

(b) reduce the data rate from kf_S back to f_S, a process known as *decimation*. Though digital filters/decimators are beyond the scope of this book, it must be said that they can be designed for very sharp cutoff characteristics with good phase response. Moreover, they are far more easily implemented and maintained with temperature and time than their analog counterparts, and they can readily be reprogrammed in the software, if needed.

We observe that the rms noise at the output of the digitizer is still $V_{FSR}/2^n\sqrt{12}$; however, only the shaded portion will make it past the filter/decimator, so the rms noise at the output is

$$E_q = \left(\int_0^{f_S/2} \frac{q^2}{kf_S/2}\,df\right)^{1/2} = q/\sqrt{k} \qquad (12.24)$$

or $E_q = V_{FSR}/2^n\sqrt{12k}$. Expressing k in the form $k = 2^m$, we now have

$$\text{SNR}_{\max} = 6.02(n + 0.5m) + 1.76 \text{ dB} \qquad (12.25)$$

indicating a $\frac{1}{2}$-bit improvement for every octave of oversampling.

EXAMPLE 12.6. An audio signal is oversampled with a 12-bit ADC. Find the oversampling frequency needed to achieve a 16-bit resolution. What is the corresponding $\text{SNR}_{\max}$?

Solution. To gain $16 - 12 = 4$ bits of resolution, we need to oversample by $m = 4/(1/2) = 8$ octaves, so the oversampling frequency must be $2^8 \times 44.1$ kHz $=$ 11.29 MHz. Moreover, $\text{SNR}_{\max} = 98.09$ dB.

Remark. Oversampling, while increasing resolution, does not improve linearity: the integral nonlinearity of the final 16-bit conversion cannot be better than that of the 12-bit ADC used!

Noise Shaping and Σ-Δ Converters

It is instructive to develop an intuitive feel for quantization-noise reduction. To this end, refer back to the 3-bit ADC example of Fig. 12.5, and suppose we apply a constant input V_I lying somewhere between $\frac{3}{8}$ V and $\frac{4}{8}$ V. The ADC will yield either $D_O = 011$ or $D_O = 100$, depending on whether V_I is closer to $\frac{3}{8}$ V or to $\frac{4}{8}$ V. Moreover, only *one sample* needs be taken to find D_O. An ingenious way to increase resolution above 3 bits is to add a Gaussian-noise dither $e_n(t)$ to V_I, and take *multiple samples* of the resulting signal $v_I(t) = V_I + e_n(t)$. Because of the fluctuations of $v_I(t)$, the samples will form a Gaussian distribution about some mean value, which we can easily compute by taking the average of our multiple readings. The result gives a more accurate estimate of V_I! In fact, Eq. (12.25) indicates that we need four samples to increase resolution by 1 bit, sixteen samples to increase by 2 bits, sixty-four samples to increase by 3 bits, and so forth.

Σ-Δ ADCs use feedback for the double purpose of (a) generating dither to keep the input busy, and (b) reshaping the noise spectrum to reduce the amount of oversampling required. In its simplest form[1] depicted in Fig. 12.38a, a Σ-Δ ADC consists of a 1-bit digitizer or modulator to convert v_I to a high-frequency serial data

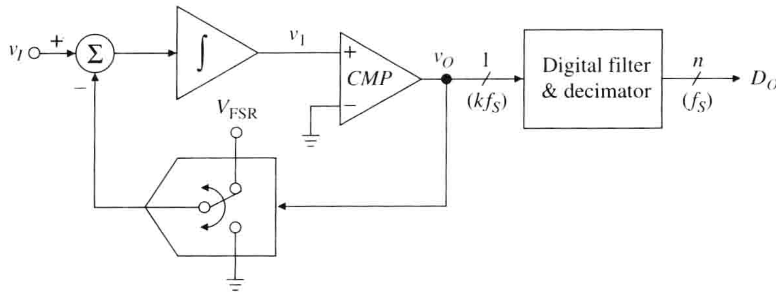


FIGURE 12.38
First-order Σ-Δ ADC.

stream v_O, followed by a digital filter/decimator to convert this stream to a sequence of n-bit words of fractional binary value D_O at a lower rate of f_S words per second. The modulator is made up of a latched comparator acting as a 1-bit ADC, a 1-bit DAC, and an integrator to integrate (Σ) the difference (Δ) between v_I and the DAC output; hence the name Σ-Δ ADC. The comparator is strobed at a rate of kf_S sps, where k, usually a power of 2, is called the *oversampling ratio.*

Figure 12.39 shows the integrator and comparator outputs for two representative input conditions (the dots mark the instants in which *CMP* is strobed). In (*a*) v_I is set at midrange, so the serial stream contains an equal number of 0s and 1s. To decode this stream with a 2-bit resolution, we pass it through a digital filter which computes its average over four samples. The result is the fractional binary value $D_O = 10$, corresponding to $(\frac{1}{2} + \frac{0}{4})V_{FSR}$, or $0.5V_{FSR}$. In (*b*) v_I is set at $\frac{3}{4}$ of the range, so the serial stream contains three 1s for every 0. After averaging, this gives $D_O = 11$, corresponding to $(\frac{1}{2} + \frac{1}{4})V_{FSR}$, or $0.75V_{FSR}$. It is apparent that the distribution of 0s and 1s in the serial stream depends on the value of v_I within the range of 0 to V_{FSR}.

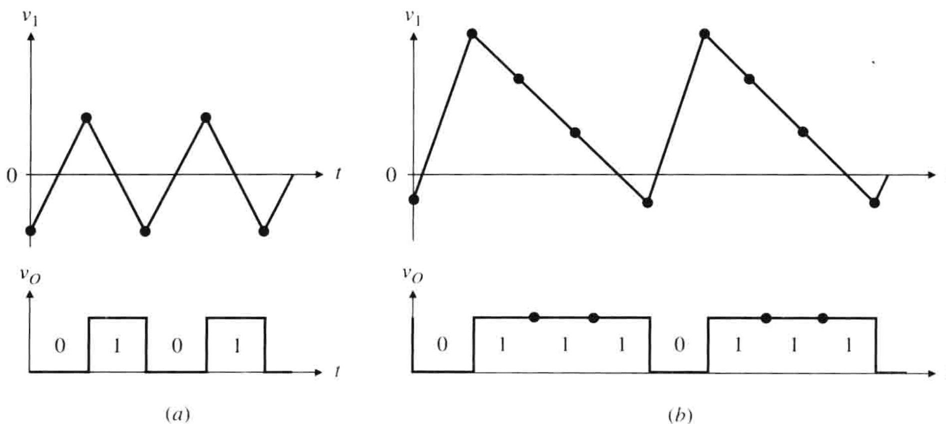


FIGURE 12.39
Integrator and comparator outputs for (*a*) $v_I = 0.5V_{FSR}$, and (*b*) $v_I = 0.75V_{FSR}$.

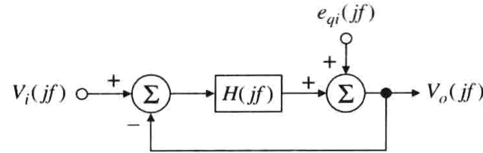


FIGURE 12.40
Linear system model of a Σ-Δ ADC.

To understand how noise shaping comes about, refer to Fig. 12.40, where the quantization error is modeled additively via the noise process $e_{qi}(jf) = q/\sqrt{kf_S/2}$. By inspection, the various Fourier transforms are related as $V_o = e_q + H \times (V_i - V_o)$, or

$$V_o(jf) = \frac{1}{1 + 1/H(jf)} V_i(jf) + \frac{1}{1 + H(jf)} e_{qi}(jf) \tag{12.26}$$

Choosing $H(jf)$ such that its magnitude is sufficiently large over the frequency band of interest will provide the simultaneous benefits of (a) making V_o *closely track* V_i over the given band and (b) *drastically reducing* quantization noise over the same band. This is not surprising for the observant reader who has already noted the similarity of Fig. 12.40 to Fig. 1.27, with H playing the role of the loop gain and with e_{qi} playing the role of x_3.

For frequency bands extending down to dc, $H(jf)$ is usually implemented with integrators; however, depending on the application, other filter types may be more efficient, such as band-pass filters in telecommunications.[11] In mixed-mode IC processes, $H(jf)$ is implemented using switched-capacitor techniques. Figure 12.41 shows an SC realization[11] of the 1-bit modulator. Using Eq. (4.22) with $C_1 = C_2$, $\omega = 2\pi f$, and $T_{CK} = 1/(kf_S)$, we can express the SC integrator transfer function as $H(jf) = 1/\{\exp[j2\pi f/(kf_S)] - 1\}$. Substituting into Eq. (12.26) gives

$$V_o(jf) = V_i(jf)e^{-j2\pi f/(kf_S)} + e_{qo}(jf) \tag{12.27}$$

$$e_{qo}(jf) = (1 - e^{-j2\pi f/(kf_S)})e_{qi}(jf) \tag{12.28}$$

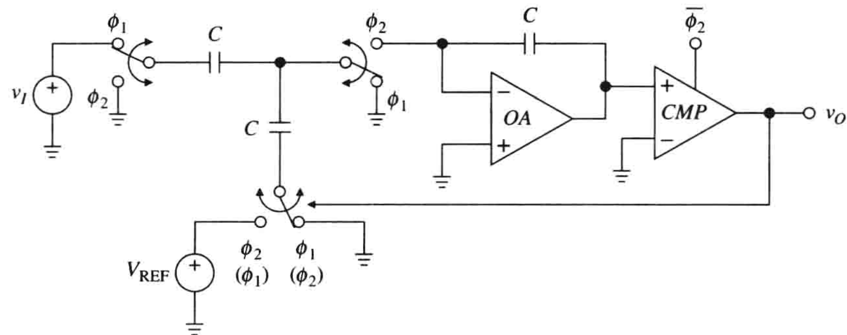


FIGURE 12.41
Switched-capacitor implementation of a first-order modulator. Bottom-switch phase is (ϕ_1, ϕ_2) for v_O = high, and (ϕ_2, ϕ_1) for v_O = low.

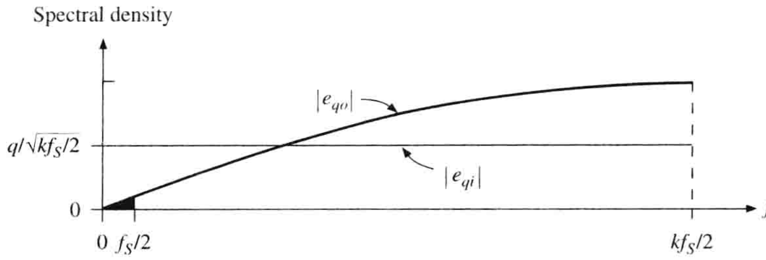


FIGURE 12.42
First-order noise shaping ($k = 16$).

By the well-known Fourier-transform property that multiplying by $\exp(-j\omega T)$ in the frequency domain is equivalent to delaying by T in the time domain, Eq. (12.27) indicates that v_O is simply v_I delayed by $1/(kf_S)$. Moreover, applying Euler's identity to Eq. (12.28), we can write

$$|e_{qo}(jf)| = 2\sin[\pi f/(kf_S)] \times |e_{qi}(jf)| \tag{12.29}$$

The plot of Fig. 12.42 reveals that the modulator shifts most of the noise energy toward higher frequencies. Only the shaded portion will make it past the filter/decimator, so the corresponding rms output noise is obtained as

$$E_q = \left(\int_0^{f_S/2} |e_{qo}(jf)|^2 df \right)^{1/2} \tag{12.30}$$

For $k \gg \pi$, we obtain (see Problem 12.22) $E_q = \pi q/\sqrt{3k^3} = \pi V_{FSR}/(2^n\sqrt{36k^3})$. Expressing k in the form $k = 2^m$ gives, for a first-order Σ-Δ ADC,

$$\mathrm{SNR}_{\max} = 6.02(n + 1.5m) - 3.41\,\mathrm{dB} \tag{12.31}$$

indicating a 1.5-bit improvement for every octave of oversampling; this is better than the 0.5-bit improvement without noise shaping.

The benefits of noise shaping can be enhanced further by using higher-order modulators. For instance, suitably cascading[11] two subtractor-integrator blocks gives a second-order Σ-Δ ADC with

$$|e_{qo}(jf)| = \{2\sin[\pi f/(kf_S)]\}^2 \times |e_{qi}(jf)| \tag{12.32}$$

Substituting into Eq. (12.30), we obtain (see Problem 12.22), for $k \gg \pi$, $E_q = \pi^2 q/\sqrt{5k^5} = \pi^2 V_{FSR}/(2^n\sqrt{60k^5})$. This yields, for a second-order Σ-Δ ADC,

$$\mathrm{SNR}_{\max} = 6.02(n + 2.5m) - 11.14\,\mathrm{dB} \tag{12.33}$$

indicating a 2.5-bit improvement for every octave of oversampling.

EXAMPLE 12.7. Find k for $\mathrm{SNR}_{\max} \geq 96$ dB (or ≥ 16 bits) using (*a*) a first-order and (*b*) a second-order Σ-Δ ADC.

Solution.

(*a*) Imposing $6.02(1 + 1.5m) - 3.41 \geq 96$ gives $m \geq 10.3$, or $k \geq 2^{10.3} \cong 1261$.
(*b*) Similarly, $k \geq 2^{6.7} \cong 105$.

Besides offering the aforementioned advantages of undemanding and mixed-mode-compatible analog circuitry, 1-bit quantizers are inherently linear: since only two output levels are provided, a straight characteristic results, with no need for trimming or calibration as in multilevel quantizers. Moreover, the presence of the integrator makes the input SHA unnecessary—if at the price of more stringent input-drive requirements due to charge injection effects.[12]

Practical upper limits on sampling rates currently restrict Σ-Δ ADCs to moderate-speed but high-resolution applications, such as digital audio, digital telephony, and low-frequency measurement instrumentation, with resolutions ranging from 16 to 24 bits.[12–14] An additional factor to keep in mind is that since the digital filter/decimator computes each high-resolution sample using many previous low-resolution samples, there is a *latency* as information progresses from input to output through the various stages of the filter. This delay may be intolerable in certain real-time applications, such as control. Moreover, it makes Σ-Δ converters unsuited to input multiplexing, that is, to situations where it is desired to share the same ADC among different sources to help reduce cost.

The interested reader is referred to the literature[9–11] for additional practical issues such as stability and idle tones, system architectures, and the fascinating subject of digital filtering and decimation.

PROBLEMS

12.1 Performance specifications

12.1 A 3-bit DAC designed for $V_{FSR} = 3.2$ V is sequenced through all input codes from 000 to 111, and the actual output values are found to be $v_O = 0.2, 0.5, 1.1, 1.4, 1.7, 2.0,$ 2.6, and 2.9, all in V. Find the offset error, the gain error, the INL, and the DNL, in fractions of 1 LSB.

12.2 A full-scale sinusoid is applied to a 12-bit ADC. If the digital analysis of the output reveals that the fundamental has a normalized power of 1 W while the remaining power is 0.6 μW, find the effective number of bits of this ADC. What is the SNR if the input sinusoid is reduced to 1/100th of full scale?

12.2 D-A conversion techniques

12.3 A 6-bit weighted-resistor DAC of the type of Fig. 12.7 is implemented with $V_{REF} =$ 1.600 V, but with $R_f = 0.99R$ instead of $R_f = R$, and a low-quality op amp having $V_{OS} = 5$ mV and $a = 200$ V/V. Find the offset and gain errors of this DAC, in fractions of 1 LSB. What is the worst-case value of the output when all bits are set to 1?

12.4 A 4-bit weighted-resistor DAC of the type of Fig. 12.7 is implemented with $V_{REF} =$ -3.200 V and a high-quality op amp, but gross resistor values, namely, $R_f = 9.0$ kΩ instead of 10 kΩ, $2R = 22$ kΩ instead of 20 kΩ, $4R = 35$ kΩ instead of 40 kΩ, $8R = 50$ kΩ instead of 80 kΩ, and $16R = 250$ kΩ instead of 160 kΩ. Find the gain error, along with the integral and differential nonlinearities. Comment on your findings.

12.5 The AH5010 quad switch consists of four analog-ground p-FET switches and relative diode clamps of the type of Fig. 9.37, plus a fifth dummy FET for $r_{ds(\text{on})}$ compensation. (a) Using an LM385 2.5-V reference diode, an AH5010 quad switch ($r_{ds(\text{on})} \cong 100\ \Omega$), and a JFET-input op amp with ±15-V supplies, design a 4-bit weighted-resistor DAC with $V_{\text{FSR}} = +10.0$ V. (b) Compute v_O for each input code. (c) Repeat if the op amp has $V_{OS} = 1$ mV. What are the offset and gain errors of your DAC?

12.6 One way of curbing excessive resistance spread in an 8-bit weighted-resistor DAC is by combining the outputs of two 4-bit DACs as $v_O = v_{O(\text{MS})} + 2^{-4}v_{O(\text{LS})}$, where $v_{O(\text{MS})}$ is the output of the DAC using the four MSBs of the 8-bit code, and $v_{O(\text{LS})}$ that of the DAC using the four LSBs. Using components of the type of Problem 12.5, design one such 8-bit DAC.

12.7 (a) Using an 8-bit R-$2R$ ladder with $R = 10\ \text{k}\Omega$, an LM385 2.5-V reference diode, and a 741 op amp, design an 8-bit voltage-mode DAC with $V_{\text{FSR}} = 10$ V. (b) Modify your circuit so that v_O is offset by −5 V. Assume ±15-V regulated supplies.

12.8 (a) Derive expressions for the element values in the Norton equivalent of the current-driven R-$2R$ ladder DAC of Fig. 12.17. (b) Suppose $V_{\text{REF}}/R_r = 1$ mA, $R = 1\ \text{k}\Omega$, and the output of the DAC is fed to a simple I-V converter op amp with a feedback resistance of 1 kΩ. If the I-V converter introduces an offset error of $\frac{1}{4}$ LSB and a gain error of $-\frac{1}{2}$ LSB, find the I-V converter output for $b_1b_2b_3b_4 = 0000$, 0100, 1000, 1100, and 1111. (c) Find the closed-loop small-signal bandwidth if the op amp has a constant GBP of 50 MHz.

12.3 Multiplying DAC applications

12.9 The programmable attenuator of Fig. 12.23a can be turned into a programmable attenuator/amplifier by using a T-network of the type of Fig. 2.2 in the feedback path. This is achieved by interposing a voltage divider between the op amp output and the R_f pin of the DAC (see Analog Devices Application Note AN-137). Using a 12-bit MDAC with $R_f = 10\ \text{k}\Omega$, design a circuit whose gain can be varied from $\frac{1}{64}$ V/V to 64 V/V as the input code is sequenced from 0...01 to 1...11.

12.10 Consider the circuit obtained from the biquad filter of Fig. 3.36 by replacing the inverting amplifier (OA_3 plus the R_3 resistances) with the programmable attenuator of Fig. 12.23a. Find an expression for the band-pass response, and verify that both f_0 and Q are proportional to $\sqrt{D}$, indicating a digitally programmable, constant-bandwidth band-pass filter.

12.11 Consider the circuit obtained from Fig. 12.24 by removing R_4, $MDAC_2$, OA_5, and the OA_3 integrator. (a) Sketch the reduced circuit, and show that now OA_1 and OA_2 provide, respectively, the first-order high-pass and low-pass responses. (b) Specify suitable components so that the low-pass response has a dc gain of 20 dB, the high-pass response has a high-frequency gain of 0 dB, and the characteristic frequency is digitally programmable in 5-Hz steps by means of a dual 10-bit MDAC.

12.12 Modify the quadrature oscillator of Fig. 10.6a for peak amplitudes of 5 V and f_0 digitally programmable in 10-Hz steps by means of a dual 10-bit MDAC.

12.13 Using a 12-bit MDAC and an AD537 wide-sweep CCO (see Fig. 10.33), design a triangular wave generator with peak values of ±5 V and f_0 digitally programmable in 10-Hz steps. The circuit is to have provision for both frequency and amplitude calibration. Assume the triangular wave available across the timing capacitor of the AD537 has a peak-to-peak amplitude of $\frac{5}{3}$ V.

12.14 Using an 8-bit CMOS DAC of the type of Fig. 12.11, an LM385 2.5-V reference diode, and an LM317 regulator of the type of Fig. 11.19, along with other components as needed, design a 1-A power supply digitally programmable over the range 0.0 V to 10.0 V. Assume ±15-V supplies.

12.4 A-D conversion techniques

12.15 As we know, an SA ADC must usually be preceded by a THA. However, if the input is sufficiently slow to change by less than $\pm\frac{1}{2}$ LSB during the conversion cycle, then the THA is unnecessary. (*a*) Show that a full-scale sine wave input can be converted without the need for a THA, provided its frequency is below $f_{max} = 1/(2^n \pi t_{SAC})$, where t_{SAC} is the time it takes for the SA ADC to complete a conversion. (*b*) Find f_{max} for an 8-bit SA ADC operating at the rate of 10^6 conversions per second. How does f_{max} change if the SA ADC is preceded by an ideal SHA?

12.16 Discuss the general requirements on the reference, DAC, and comparator of an 8-bit SA ADC for a conversion time of 1 μs over the range $0\ °C \le T \le 50\ °C$ with an accuracy of $\pm 1/2$ LSB, if $V_{FSR} = 10$ V.

12.17 Consider a charge-redistribution ADC of the type of Fig. 12.29 with $n = 4$, $V_{REF} = 3.0$ V, and $C = 8$ pF. Assuming node v_P has a parasitic capacitance of 4 pF toward ground, find the intermediate values of v_P during the conversion of $v_I = 1.00$ V.

12.18 Assume the 8-bit subranging ADC of Fig. 12.32 has $V_{REF} = 2.560$ V. (*a*) Find the total number of comparators, their voltage reference levels, and the maximum level tolerances allowed for a $\pm\frac{1}{2}$ LSB accuracy. (*b*) Find $b_1 \ldots b_8$, v_{RES}, and the quantization error for $v_I = 0.5$ V, 1.054 V, and 2.543 V.

12.19 Show that if the input to the dual-slope ADC of Fig. 12.34 contains an unwanted ac component of the type $v_i = V_m \cos(\omega t + \theta)$, then the result of integrating it over the interval $T = 2^n T_{CK}$ is proportional to the *sampling function* $\mathrm{Sa}(\omega T) = \sin(\omega T)/(\omega T)$. Plot $|\mathrm{Sa}(\omega T)|_{dB}$ vs. ωT, and verify that this type of ADC inherently rejects all unwanted ac components whose frequencies are integral multiples of $1/T$.

12.20 The integrator of a dual-slope ADC is implemented with an op amp having gain $a = 10^3$ V/V. (*a*) Assuming its output $v_O(t)$ is initially zero, find $v_O(t \ge 0)$ if the input is $v_I = 1$ V. (*b*) Find the minimum value of *RC* so that $v_O(t = 100\text{ ms})$ is afflicted by an error of less than 1 mV.

12.21 A 14-bit dual-slope ADC of the type of Fig. 12.34 is to be designed so that it rejects the 60-Hz power-line interference frequency and harmonics thereof. (*a*) What is the required clock frequency f_{CK}? What is the time required to convert a full-scale input? (*b*) If $V_{REF} = 2.5$ V and the input is in the range 0 to 5 V, what is the value of *RC* for a peak value of 5 V at the integrator's output for a full-scale input?

(c) If component aging causes R to change by $+5\%$ and C by -2%, what is the effect upon the integrator's output for the case of a full-scale input? Upon the conversion's accuracy?

12.5 Oversampling converters

12.22 (a) Plot $|e_{qo}(jf)|, 0 \leq f \leq kf_S/2$ for the second-order Σ-Δ ADC, and compare with the first order. (b) Show that the rms noise before digital filtering is $\sqrt{2}q$ for the first-order modulator, and $\sqrt{8}q$ for the second-order modulator. (c) Using the approximation $\sin x \cong x$ for $x \ll 1$, show that the rms noise after digital filtering is, for $k \gg \pi$, $\pi q/\sqrt{3k^3}$ for the first-order modulator, and $\pi^2 q/\sqrt{5k^5}$ for the second-order modulator. (d) Find the rms noise percentage removed by the digital filter for both orders if $k = 16$.

12.23 Compare the sampling rates needed for a 16-bit audio ADC using a 1-bit ADC with (a) straight oversampling, (b) first-order noise shaping, and (c) second-order noise shaping.

12.24 An 8-bit ADC that is linear to 12 bits is used to perform conversions over a 100-kHz signal bandwidth. (a) Find the sampling rate required to achieve 12 bits of accuracy using straight oversampling. (b) Repeat, but for the case in which the above ADC is placed inside a first-order Σ-Δ modulator. (c) Repeat, but for a second-order modulator.

12.25 An oversampling audio ADC with $n = 16$, $V_{FSR} = 2$ V, $f_S = 48$ kHz, and $kf_S = 64f_S$ uses a simple RC network as the input antialiasing filter. (a) Specify RC for a maximum attenuation of 0.1 dB for $0 \leq f \leq 20$ kHz. (b) Assuming the spectral makeup of v_I within the first image band $kf_S \pm 20$ kHz is just white noise with spectral density e_{nw}, find the maximum allowed value of e_{nw} so that the corresponding base-band rms noise is less than $\frac{1}{2}$ LSB.

REFERENCES

1. Analog Devices Engineering Staff, *Mixed-Signal Design Seminar,* Analog Devices, Norwood, MA, 1991.
2. M. Oljaca and B. Baker, "Start with the Right Op Amp When Driving SAR ADCs," *EDN*, Oct. 16, 2008, http://www.edn.com/contents/images/6602451.pdf.
3. Analog Devices Engineering Staff, *Analog-Digital Conversion Handbook,* 3d ed., Prentice Hall, Englewood Cliffs, NJ, 1986.
4. B. Razavi, *Principles of Data Conversion System Design,* IEEE Press, Piscataway, NJ, 1995.
5. A. P. Brokaw, "Analog Signal-Handling for High Speed and Accuracy," Analog Devices Application Note AN-342, http://www.analog.com/static/imported-files/application_notes/29047480461674310\66AN-342.pdf.
6. J. Wilson, G. Whitmore, and D. Sheingold, *Application Guide to CMOS Multiplying D/A Converters,* Analog Devices, Norwood, MA, 1978.
7. J. Williams, "Build Your Own A/D Converter for Optimum Performance," *EDN,* Mar. 20, 1986, pp. 191–198.
8. P. Klonowski, "Analog-to-Digital Conversion Using Voltage-to-Frequency Converters," Analog Devices Application Note AN-276, http://www.analog.com/static/imported-files/application_notes/185321581AN276.pdf.

9. M. W. Hauser, "Principles of Oversampling A/D Conversion," *J. Audio Eng. Soc.*, January/February 1991, pp. 3–26.
10. J. C. Candy and G. C. Temes, editors, *Oversampling Delta-Sigma Data Converters,* IEEE Press, Piscataway, NJ, 1992.
11. D. A. Johns and K. W. Martin, *Analog Integrated Circuit Design,* John Wiley & Sons, New York, 1997.
12. W. Kester, *Practical Analog Design Techniques*, Analog Devices, Norwood, MA, 1995. ISBN 0-916550-16-8.
13. Burr-Brown Engineering Staff, *Burr-Brown Design Seminar Manual*, 1996, http://www.datasheetcatalog.com/datasheets_pdf/D/E/S/I/DESIGN_SEMINAR.shtml.
14. W. Kester, *High Speed Design Techniques*, Analog Devices, Norwood, MA, 1996. ISBN-0-916550-17-6.

13

NONLINEAR AMPLIFIERS AND PHASE-LOCKED LOOPS

双极型晶体管高度可预测的特性可以被用来实现一些非常有用的非线性功能，例如对数转换和可变跨导乘法。这些功能反过来恰恰又是许多其他模拟运算的基础，例如反对数放大、有效值换算、模拟除法和开方计算、多种形式的线性化、压控放大器、滤波器和振荡器等。这些标准模块大大简化了模拟设计，同时拓宽了实际模拟电路的应用范围，尤其在考虑速度与成本以后，要求用模拟而不是数字形式的场合。

本章重点

本章从对数 / 反对数放大器开始，重点讨论相关的问题（例如稳定性和温度补偿）。接着介绍模拟乘法器及其他在模拟运算中的应用。本章进一步介绍运算跨导放大器（OTA）和它在实际模块中的使用，这些模块包括 *gm*-C 滤波

器、大动态范围电压 / 电流控制放大器、振荡器和滤波器等模块。最后以锁相环（PLL）结束。从某种程度上来说，锁相环集成了许多在前面几章中提到的要点，例如：振荡器、滤波器和频率补偿。使用常见的 4046 单片 PLL 作为例子来解释概念和应用。

13.1 LOG/ANTILOG AMPLIFIERS

A logarithmic amplifier—also called *log amp,* or *logger*—is an *I-V* converter with a transfer characteristic of the type

$$v_O = V_o \log_b \frac{i_I}{I_i} \tag{13.1}$$

where V_o is called the *output scale factor,* I_i the *input reference current,* and b is the *base,* usually 10 or 2. V_o represents the sensitivity of the log amp, in volts per decade (or per octave), and I_i is the value of i_I for which $v_O = 0$. Note that for proper operation we must always have $i_I/I_i > 0$. The quantity

$$\text{DR} = \log_b \frac{|i_I|_{\max}}{|i_I|_{\min}} \tag{13.2}$$

is called the *dynamic range* and is expressed in decades or in octaves, depending on b. For instance, a logger designed to operate over the range 1 nA $\le i_I \le$ 1 mA has DR $= \log_{10}(10^{-3}/10^{-9}) = 6$ decades, or DR $= \log_2 10^6 \cong 20$ octaves.

Plotting Eq. (13.1) on semilog paper with i_I/I_i on the logarithmic axis and v_O on the linear axis, as in Fig. 13.1*a*, yields a straight line with a slope of V_o V/dec. Any departure of the actual characteristic from the best-fit straight line is called the *log conformity error* e_O. Though this error can only be observed at the output, it is convenient to refer it to the input because of the unique log-function property that equal percentage errors at the input produce equal incremental errors at the output, regardless of the point on the curve. Indeed, denoting the percentage input error as p, we have $e_O = v_{O(\text{actual})} - v_{O(\text{ideal})} = V_o \log_b[(1+p)(i_I/I_i)] - V_o \log_b[i_I/I_i]$, or

$$e_O = V_o \log_b(1+p) \tag{13.3}$$

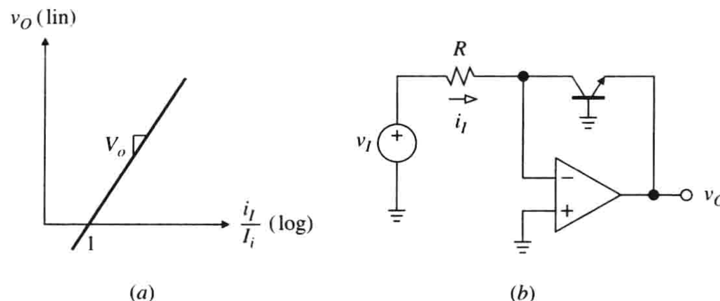


FIGURE 13.1
Logarithmic characteristic and the transdiode configuration.

For instance, with $b = 10$ and $V_o = 1$ V/dec, a 1% input error corresponds to an output error $e_O = 1 \log_{10}(1 + 0.01) = 4.32$ mV. Conversely, $e_O = 10$ mV corresponds to a percentage error p such that $10 \text{ mV} = 1 \log_{10}(1 + p)$, or $p = 2.33\%$.

The main application of log amps is data compression. As an example, consider the digitization of a photodetector current over the range $10 \text{ nA} \le i_I \le 100\ \mu\text{A}$ with an error of less than 1% of its actual value. Since we have a four-decade range, the required resolution is $0.01/10^4 = 1/10^6$, or 1 ppm. Since $10^6 \cong 2^{20}$, this requires a 20-bit A/D converter, which can be a challenging and expensive proposition. Consider now the effect of compressing the input with a log amp before digitizing. Letting, for instance, $b = 10$, $V_o = 1$ V/dec, and $I_i = 10$ nA, the current range is now compressed to a 0 to 4-V voltage range. Since a 1% current accuracy corresponds to a 4.32-mV voltage interval, the required resolution is now $(4.32 \times 10^{-3})/4 \cong 1/926 \cong 1/2^{10}$, or 10 bits. This represents a substantial reduction in cost and circuit complexity!

The inverse function of logarithmic compression is exponential expansion. This is provided by the *antilogarithmic amplifier* (antilog amp), whose transfer characteristic is

$$i_O = I_o b^{v_I/V_i} \tag{13.4}$$

where I_o is the *output reference current* and V_i the *input scale factor,* in volts per decade or per octave. The output of an antilog amp can be converted to a voltage by means of an op amp *I-V* converter. When plotted on semilog paper with v_I on the linear axis and i_O/I_o on the logarithmic axis, Eq. (13.4) also yields a straight line. The above log conformity error considerations still hold, but with the input and output errors interchanged.

The Transdiode Configuration

Log/antilog amplifiers exploit the exponential characteristic of a forward-active BJT. By Eq. (5.3), this characteristic can be written as $v_{BE} = V_T \ln(i_C/I_s)$. Practical logging BJTs conform to this equation remarkably well over a range of at least six decades,[1] typically for $0.1 \text{ nA} \le i_C \le 0.1$ mA. The heart of log/antilog amps is the circuit of Fig. 13.1*b*, known as the *transdiode configuration.* The op amp converts v_I to the current $i_I = v_I/R$, and then forces the BJT in its feedback path to respond with a logarithmic base-emitter voltage drop to yield

$$v_O = -V_T \ln \frac{v_I}{RI_s} \tag{13.5}$$

If we also take into account the input bias voltage V_{OS} and bias current I_B, then the collector current becomes $i_C = (v_I - V_{OS})/R - I_B$, so the transfer characteristic takes on the more realistic form

$$v_O = -V_T \ln \frac{v_I - V_{OS} - RI_B}{RI_s} \tag{13.6}$$

The input offset error $(V_{OS} + RI_B)$ sets the ultimate limit on the range of inputs that can be processed within a given log conformity error. Wide-dynamic-range loggers use op amps with ultra-low V_{OS} and I_B to approach the ideal characteristic of Eq. (13.5). The ultimate limit is then posed by drift and noise. If the transdiode is driven directly with a current source i_I, Eq. (13.5) reduces to $v_O = -V_T \ln(i_I/I_s)$, and the ultimate limit is now set by the input bias current of the op amp or by the

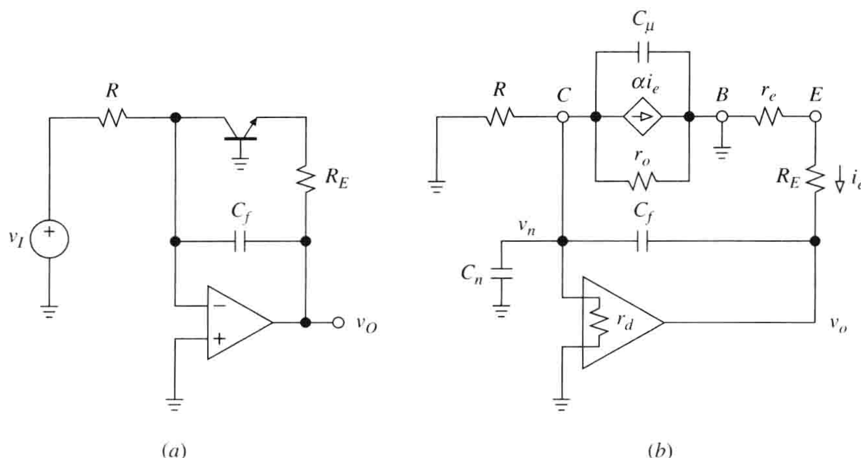


FIGURE 13.2
Transdiode circuit with frequency compensation, and its incremental model.

low-end log conformity error of the BJT, whichever is higher. In general, current-driven loggers offer a wider dynamic range than voltage-driven loggers.

Stability Considerations

Transdiode circuits are notorious for their tendency to oscillate due to the presence of an active gain element inside the feedback loop. As shown in Fig. 13.2*a*, the transdiode is stabilized[1] by using an emitter-degeneration resistance R_E to reduce the feedback factor β, and a feedback capacitance C_f to provide feedback lead. To investigate stability, we need to find the feedback factor β. To this end, refer to the ac model of Fig. 13.2*b*, where the BJT has been replaced by its common-base small-signal model.[2] The BJT parameters r_e and r_o depend on the operating current I_C as

$$r_e = \frac{\alpha V_T}{I_C} \qquad r_o = \frac{V_A}{I_C} \tag{13.7}$$

where V_T is the thermal voltage and V_A is the so-called Early voltage. Typically, $\alpha \cong 1$ and $V_A \cong 100$ V. The base-collector junction capacitance C_μ and the inverting-input stray capacitance C_n are typically on the order of a few picofarads.

Circuit analysis is facilitated by the introduction of

$$R_a = R \parallel r_o \parallel r_d \qquad R_b = r_e + R_E \tag{13.8}$$

Applying KCL at the summing junction gives

$$v_n[1/R_a + j\omega(C_n + C_\mu)] + \alpha i_e + j\omega C_f(v_n - v_o) = 0$$

Letting $i_e = -v_o/R_b$, rearranging, and solving for $\beta = v_n/v_o$ gives, for $\alpha \cong 1$,

$$\frac{1}{\beta} = \frac{R_b}{R_a}\frac{1 + jf/f_z}{1 + jf/f_p} \tag{13.9}$$

where the zero and pole frequencies are

$$f_z = \frac{1}{2\pi R_a(C_n + C_\mu + C_f)} \qquad f_p = \frac{1}{2\pi R_b C_f} \tag{13.10}$$

The $|1/\beta|$ curve has the low-frequency asymptote $1/\beta_0 = R_b/R_a$, the high-frequency asymptote $1/\beta_\infty = 1 + (C_n + C_\mu)/C_f$, and two breakpoints at f_z and f_p. While f_z and $1/\beta_\infty$ are relatively constant, f_p and $1/\beta_0$ depend on the operating current as per Eq. (13.7), so they can vary over a wide range of values, as exemplified in Fig. 13.3. The hardest condition to compensate is when i_C is maximized, since this minimizes $1/\beta_0$ and maximizes f_p, leading to the highest rate of closure. As a rule of thumb,[1] R_E is chosen so that, when i_C is maximized, $1/\beta_0 \cong 0.5$ V/V and $f_p \cong 0.5 f_x$, where f_x is the crossover frequency.

EXAMPLE 13.1. In the circuit of Fig. 13.2*a* let $R = 10$ kΩ, 1 mV $< v_I <$ 10 V, $C_n + C_\mu = 20$ pF, $V_A = 100$ V, $r_d = 2$ MΩ, and $f_t = 1$ MHz. Find suitable values for R_E and C_f.

Solution. At the upper end of the range, where $i_C = (10\text{ V})/(10\text{ k}\Omega) = 1$ mA, we have $r_e \cong 26\ \Omega$, $r_o = 100$ kΩ, and $R_a \cong 9$ kΩ. Imposing $(26 + R_E)/9000 = 0.5$ gives $R_E = 4.47$ kΩ (use 4.3 kΩ).

Next, find f_x using the definition $|a(jf_x)| \times \beta_\infty = 1$. Letting $|a(jf_x)| \cong f_t/f_x$ and using $1/\beta_\infty = 1 + (C_n + C_\mu)/C_f$ we get $f_x = f_t/[1 + (C_n + C_\mu)/C_f]$. Imposing $f_p = 0.5 f_x$ and simplifying, we finally obtain

$$C_f = \frac{1 + (C_n + C_\mu)/C_f}{\pi R_b f_t}$$

Substituting the given parameter values, along with $R_b \cong 4.3$ kΩ, gives $C_f = 90$ pF (use 100 pF).

Figure 13.3 indicates that at low values of v_I the response is dominated by f_p, thus resulting in slow dynamics. This is not surprising, since at low current levels it takes longer to charge or discharge the various capacitances. At low currents

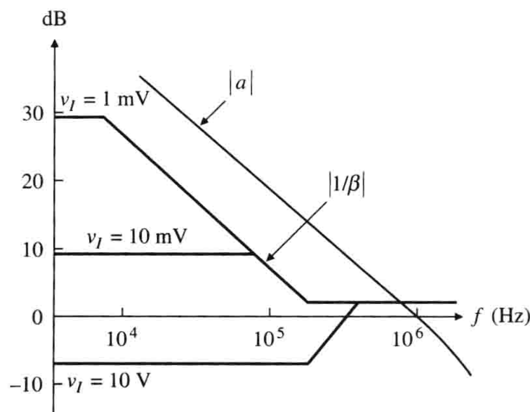


FIGURE 13.3
Bode plots for the transdiode circuit of Example 13.1.

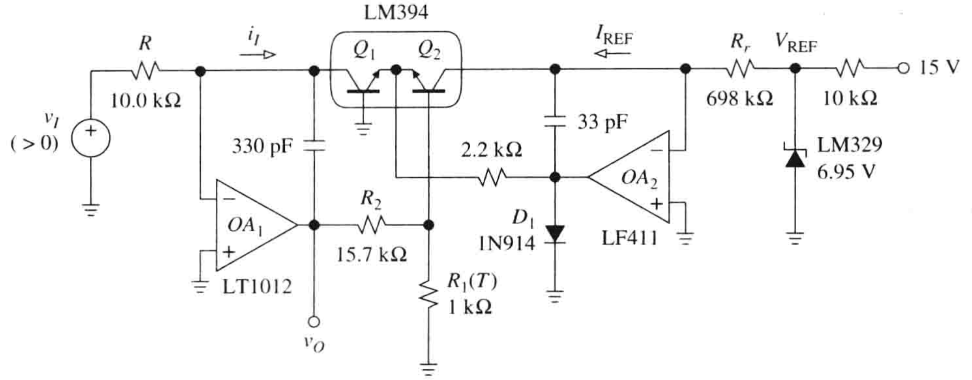


FIGURE 13.4
Logarithmic amplifier.

we have $r_e \gg R_E$, so $f_p \cong 1/(2\pi r_e C_f)$, indicating a time constant $\tau \cong r_e C_f \cong (V_T/I_C)C_f = (V_T/v_I)RC_f$. For instance, with $C_f = 100$ pF, at $I_C = 1$ nA we have $\tau \cong (0.026/10^{-9})10^{-10} = 2.6$ ms, so we must be prepared for slow dynamics near the low end of the range.

Practical Log/Antilog Circuits[3]

Both the output scale factor and the input reference term in Eq. (13.5) depend on temperature. The circuit of Fig. 13.4 overcomes this serious drawback by using a matched BJT pair to eliminate I_s, and a temperature-sensitive voltage divider to compensate for TC(V_T). The op amps force the BJTs to develop $v_{BE1} = V_T \ln(i_I/I_{s1})$ and $v_{BE2} = V_T \ln(I_{REF}/I_{s2})$, where $I_{REF} = V_{REF}/R_r$. By the voltage divider formula, $v_{B2} = v_O/(1 + R_2/R_1)$. But, by KVL, $v_{B2} = v_{BE2} - v_{BE1} = V_T \ln[(I_{REF}/I_{s2})(I_{s1}/i_I)]$. Eliminating v_{B2} and using the property $\ln x = 2.303 \log_{10} x$, we get

$$v_O = V_o \log_{10} \frac{i_I}{I_i} \tag{13.11}$$

$$V_o = -2.303 \frac{R_1(T) + R_2}{R_1(T)} V_T \qquad I_i = \frac{V_{REF}}{R_r} \times \frac{I_{s2}}{I_{s1}} \tag{13.12}$$

For matched BJTs $I_{s2}/I_{s1} = 1$, so we get the temperature-independent expression $I_i = V_{REF}/R_r$. Moreover, for $R_2 \gg R_1$, we can approximate $V_o \cong -2.303 R_2 V_T / R_1(T)$, indicating that V_o can be thermally stabilized by using a resistance $R_1(T)$ with TC(R_1) = TC(V_T) = $1/T$ = 3660 ppm/°C. A suitable resistor is the Q81 (Tel Labs), which must be mounted in close thermal coupling with the BJT pair. The function of D_1 is to protect the BJTs against inadvertent reverse bias. The use of the LT1012 picoampere-input-current, microvolt-offset, low-noise op amp allows for a voltage-logging range of $4\frac{1}{2}$ decades. With the given component values, $V_o = -1$ V/dec and $I_i = 10\ \mu$A, so $v_O = -(1 \text{ V/dec}) \log_{10}[v_I/(0.1 \text{ V})]$. V_o and I_i are calibrated via R_2 and R_r.

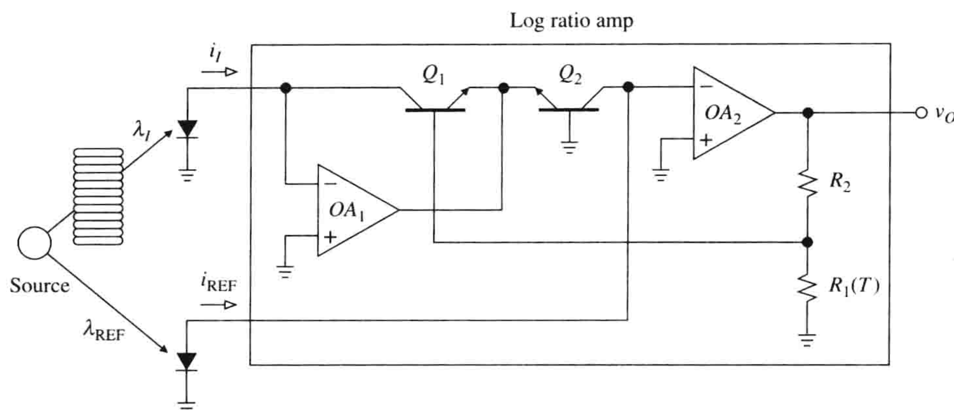


FIGURE 13.5
Log ratio amplifier for absorbance measurements.

If the input reference current I_i is allowed to vary, the log amp is called a *log ratio amplifier* and finds application in wide-dynamic-range ratiometric measurements where the unknown signal is measured against a reference signal that is itself variable. Typical examples are absorbance measurements in medicine and pollution control, where light transmitted through a specimen is measured against incident light, and the result must be independent of incident light intensity. This application is illustrated in Fig. 13.5, where frequency compensation and reverse-bias protection have been omitted for simplicity. The transmitted light λ_I and the incident light λ_{REF} are converted to the proportional currents i_I and i_{REF} by a pair of matched photodiodes operating in the photovoltaic mode. Then, the circuit computes the log ratio $v_O = V_o \log_{10}(i_I/i_{REF}) = V_o \log_{10}(\lambda_I/\lambda_{REF})$, where V_o is given by Eq. (13.12).

Figure 13.6 shows how the log amp can be rearranged to implement an exponential amp. It is left as an exercise for the reader (see Problem 13.4) to prove that the circuit gives

$$i_O = I_o 10^{v_I/V_i} \tag{13.13}$$

$$I_o = \frac{V_{REF}}{R_r} \times \frac{I_{s2}}{I_{s1}} \qquad V_i = -2.303 \frac{R_1(T) + R_2}{R_1(T)} V_T \tag{13.14}$$

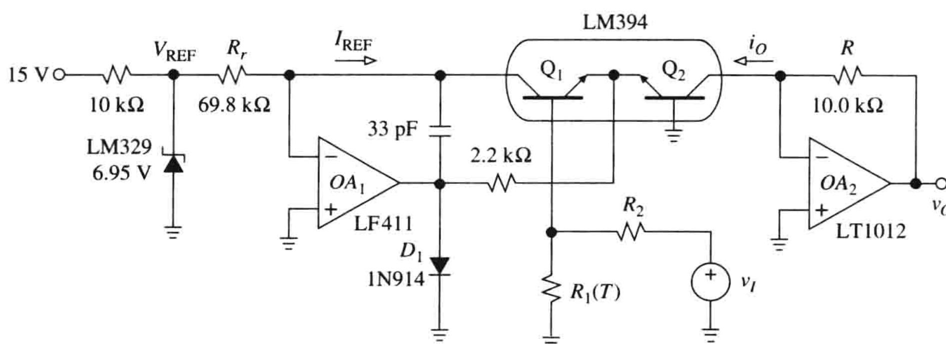


FIGURE 13.6
Antilog amplifier.

With the given component values, $I_o = 0.1$ mA and $V_i = -1$ V/dec. It is important that the collector of Q_2 be returned to a 0-V node, such as the virtual-ground node of the *I-V* converter OA_2, in order to nullify the collector-base leakage current of Q_2. Otherwise, this current may degrade log conformity at the low end of the range.

Log, log-ratio, and antilog amplifiers are available in IC form from various manufacturers. These devices usually work over a six-decade current range (1 nA to 1 mA) and a four-decade voltage range (1 mV to 10 V).

True rms-to-dc Converters

The logarithmic characteristics of BJTs are also exploited to perform a variety of slide-rule-like analog computations. A popular example is *true rms-to-dc conversion,* defined as

$$V_{\text{rms}} = \left(\frac{1}{T} \int_0^T v^2(t)\, dt \right)^{1/2} \tag{13.15}$$

V_{rms} gives a measure of the energy content of $v(t)$, so it provides the basis for accurate and consistent measurements, especially in the case of ill-defined waveforms, such as noise (electronic noise, switch contact noise, acoustical noise), mechanical transducer outputs (stress, vibration, shock, bearing noise), SCR waveforms, low-repetition-rate pulse trains, and other waveforms carrying information on the average energy generated, transmitted, or dissipated.

Equation (13.15) can be mechanized by performing the operations of squaring, averaging, and square rooting. Referred to as *explicit* rms computation, this scheme places severe demands on the dynamic output range of the squarer, which must be twice as wide as the input range. This drawback is overcome by *implicit* rms computation, in which the gain of the squarer is made inversely proportional to V_{rms} to make the output dynamic range comparable to the input range.

A common implementation of this principle is shown in Fig. 13.7, where frequency compensation and reverse-bias protection have been omitted for simplicity. OA_1 and the associated circuitry convert $v(t)$ to a full-wave rectified current $i_{C1} = |v|/R$ flowing into Q_1. By KVL, $v_{BE3} + v_{BE4} = v_{BE1} + v_{BE2}$, or $V_T \ln[(i_{C3}/I_{s3}) \times (i_{C4}/I_{s4})] = V_T \ln[(i_{C1}/I_{s1}) \times (i_{C2}/I_{s2})]$. Assuming pairwise matched BJTs and ignoring base currents so that $i_{C2} = i_{C1}$, we get

$$i_{C3} = \frac{i_{C1}^2}{i_{C4}}$$

Substituting $i_{C1}^2 = v^2/R^2$ and $i_{C4} = V_{\text{rms}}/R$ gives

$$i_{C3} = \frac{v^2}{R V_{\text{rms}}}$$

indicating that the scale factor of the squaring function is controlled by V_{rms}, as expected of implicit computation.

OA_3 forms a low-pass filter with cutoff frequency $f_0 = 1/(2\pi RC)$. For signal frequencies sufficiently higher than f_0, OA_3 will provide the running average of

FIGURE 13.7
True rms converter.

i_{C3} as $V_{rms} \cong \overline{Ri_{C3}} = \overline{v^2/V_{rms}}$. Making the approximation $\overline{V_{rms}} \cong V_{rms}$, we can write $V_{rms}^2 = \overline{v^2}$, or

$$V_{rms} = (\overline{v^2})^{1/2} \tag{13.16}$$

As a consequence of the approximations made, V_{rms} of Eq. (13.16) will differ from the ideal V_{rms} of Eq. (13.15) by an average (or dc) error as well as an ac (or ripple) error. Both errors can be kept below a specified limit by using a suitably large capacitance.[4] However, too large a capacitance will increase the response time of the circuit, so a compromise must be reached. An effective way of reducing ripple without unduly lengthening the response is to use a *post filter,* such as a low-pass *KRC* type.

The structure of Fig. 13.7 (or improved variations thereof) is available in IC form from various manufacturers. Consult the literature[4] for useful application tips.

13.2
ANALOG MULTIPLIERS

A multiplier produces an output v_O proportional to the product of two inputs v_X and v_Y,

$$v_O = kv_X v_Y \tag{13.17}$$

where k is a scale factor, usually $1/10\ \text{V}^{-1}$. A multiplier that accepts inputs of either polarity and preserves the correct polarity relationship at the output is referred to as a *four-quadrant multiplier*. Both the input and output ranges are usually from -10 V to $+10$ V. By contrast, a *two-quadrant* multiplier requires that one of its inputs be unipolar, and a *one-quadrant multiplier* requires that both inputs be unipolar.

Multiplier performance is specified in terms of *accuracy* and *nonlinearity.* Accuracy represents the maximum deviation of the actual output from the ideal value predicted by Eq. (13.17); this deviation is also referred to as the *total error.* Nonlinearity, also referred to as *linearity error,* represents the maximum output deviation from the best-fit straight line for the case where one input is varied from end to end while the other is kept fixed, usually at +10 V or −10 V. Both accuracy and nonlinearity are expressed as a percentage of the full-scale output.

Multiplier dynamics are specified in terms of the *small-signal bandwidth,* representing the frequency where the output is 3 dB below its low-frequency value, and the *1% absolute-error bandwidth,* representing the frequency where the output magnitude starts to deviate from its low-frequency value by 1%.

Variable-Transconductance Multipliers

Monolithic four-quadrant multipliers utilize the *variable-transconductance principle*[5] to achieve errors of fractions of 1% over small-signal bandwidths extending well into the megahertz range. This principle is illustrated in Fig. 13.8*a*. The block uses the differential pair Q_3-Q_4 to provide variable transconductance, and the diode-connected pair Q_1-Q_2 to provide the proper base drive for the former. The following analysis assumes matched BJTs and negligible base currents.

By KVL, $v_{BE1} + v_{BE4} - v_{BE3} - v_{BE2} = 0$, or $v_{BE3} - v_{BE4} = v_{BE1} - v_{BE2}$. Using the logarithmic *v-i* characteristics of the BJTs, this can be expressed as $V_T \ln(i_3/i_4) = V_T \ln(i_1/i_2)$, or

$$\frac{i_3}{i_4} = \frac{i_1}{i_2}$$

Rewriting as $(i_3 - i_4)/(i_3 + i_4) = (i_1 - i_2)/(i_1 + i_2)$ gives

$$i_3 - i_4 = \frac{(i_1 - i_2) \times (i_3 + i_4)}{i_1 + i_2} \tag{13.18}$$

indicating the circuit's ability to multiply the current difference $(i_1 - i_2)$ by the total emitter current $(i_3 + i_4)$.

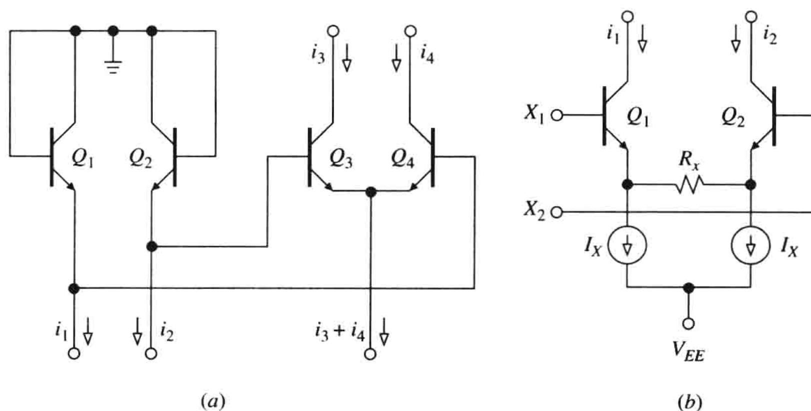


FIGURE 13.8
Linearized transconductance block and differential *V-I* converter.

To be of practical use, the circuit requires two *V-I* converters to synthesize the terms $(i_1 - i_2)$ and $(i_3 + i_4)$ from the input voltages v_X and v_Y, and an *I-V* converter to convert $(i_3 - i_4)$ to the output voltage v_O. Moreover, provisions must be made to ensure four-quadrant operation; as is, the circuit is only two-quadrant because the current $(i_3 + i_4)$ must always flow out of the emitters.

Figure 13.8*b* shows the circuit used to provide *V-I* conversion. By KCL, $i_1 = I_X + i_{R_x}$ and $i_2 = I_X - i_{R_x}$, where $i_{R_x} = (v_{E1} - v_{E2})/R_x$ is the current through R_x, assumed to flow from left to right. Consequently,

$$i_1 - i_2 = 2\frac{v_{E1} - v_{E2}}{R_x}$$

By KVL, $v_{E1} - v_{E2} = (v_{X_1} - v_{BE1}) - (v_{X_2} - v_{BE2}) = (v_{X_1} - v_{X_2}) - (v_{BE1} - v_{BE2})$, or

$$v_{E1} - v_{E2} = v_{X_1} - v_{X_2} - V_T \ln \frac{i_1}{i_2}$$

Combining the two equations gives

$$i_1 - i_2 = \frac{2}{R_x}(v_{X_1} - v_{X_2}) - \frac{2V_T}{R_x} \ln \frac{i_1}{i_2} \tag{13.19}$$

In a well-designed multiplier the last term is on the order of 1% of the other two, so we can ignore it and approximate

$$i_1 - i_2 = \frac{2}{R_x}(v_{X_1} - v_{X_2}) \tag{13.20}$$

indicating the circuit's ability to provide differential *V-I* conversion.

Figure 13.9 shows the complete multiplier. Four-quadrant operation is achieved by using two transconductance pairs with the bases driven in antiphase and the emitters driven by a second *V-I* converter. Substituting Eq. (13.20) into Eq. (13.18) and using the identities $i_1 + i_2 = 2I_X$ and $i_3 + i_4 = i_9$, we obtain

$$i_3 - i_4 = \frac{v_{X_1} - v_{X_2}}{R_x I_X} i_9$$

Likewise, using the identity $i_5 + i_6 = i_{10}$, we obtain

$$i_6 - i_5 = \frac{v_{X_1} - v_{X_2}}{R_x I_X} i_{10}$$

Subtracting the first equation from the second pairwise and using $i_{10} - i_9 = (2/R_y)(v_{Y_1} - v_{Y_2})$, we obtain

$$(i_4 + i_6) - (i_3 + i_5) = \frac{(v_{X_1} - v_{X_2})(v_{Y_1} - v_{Y_2})}{R_x R_y I_X/2}$$

The output *I-V* converter is made up of the op amp and a third *V-I* converter in its feedback path, namely, Q_{11}-Q_{12}. By KVL, the voltages at the inverting and noninverting inputs are $v_N = V_{CC} - R(i_4 + i_6 + i_{11})$ and $v_P = V_{CC} - R(i_3 + i_5 + i_{12})$. The op amp will provide Q_{12} with whatever drive it takes to make $v_N = v_P$, or $i_4 + i_6 + i_{11} = i_3 + i_5 + i_{12}$, that is,

$$(i_4 + i_6) - (i_3 + i_5) = i_{12} - i_{11} = \frac{2}{R_z}(v_{Z_1} - v_{Z_2})$$

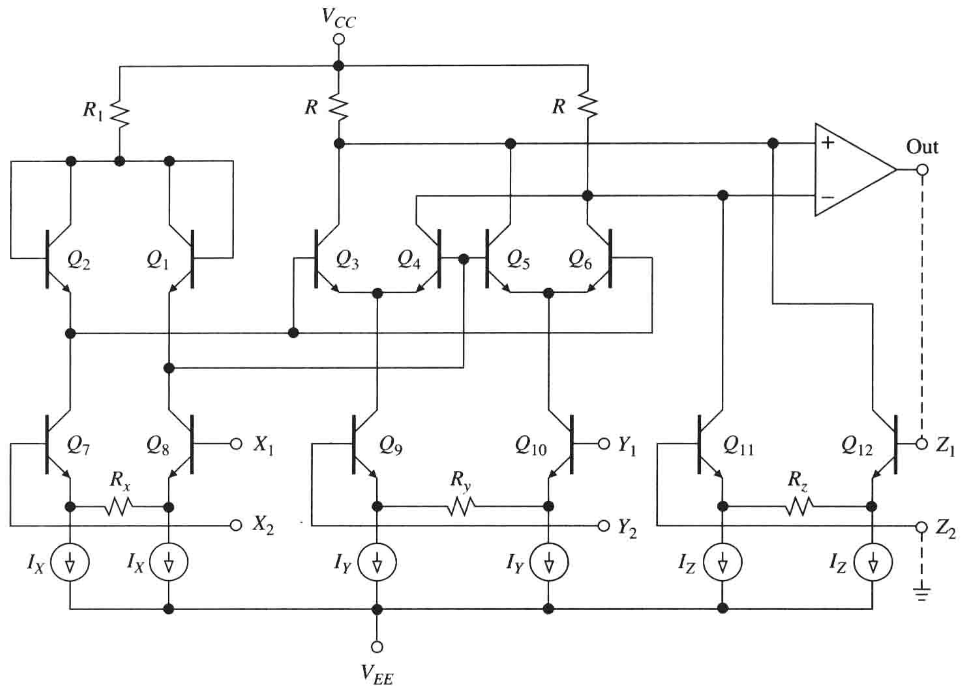


FIGURE 13.9
Four-quadrant analog multiplier.

Combining the last two equations, we finally obtain

$$v_{Z_1} - v_{Z_2} = k(v_{X_1} - v_{X_2})(v_{Y_1} - v_{Y_2}) \tag{13.21}$$

$$k = \frac{R_z}{R_x R_y I_X} \tag{13.22}$$

Most multipliers are designed for $k = 1/(10\ \text{V})$. Letting $v_O = v_{Z_1} - v_{Z_2}$, $v_X = v_{X_1} - v_{X_2}$, and $v_Y = v_{Y_1} - v_{Y_2}$ gives Eq. (13.17).

One of the main causes of linearity error is the logarithmic term of Eq. (13.19). This error is, to a first approximation, compensated for by introducing an equal but opposite nonlinearity term via the *V-I* converter Q_{11}-Q_{12} inside the feedback path. The architecture of Fig. 13.9 forms the basis of a variety of monolithic multipliers. Two of the earliest and most popular examples are the AD534 and MPY100. The AD534L version has a maximum pretrimmed total error of 0.25%, a maximum linearity error of 0.12%, a typical small-signal bandwidth of 1 MHz, and a typical 1% amplitude error bandwidth of 50 kHz.

Multiplier Applications

Analog multipliers find application in signal modulation/demodulation, analog computation, curve fitting, transducer linearization, CRT distortion compensation, and a variety of voltage-controlled functions.[1,6]

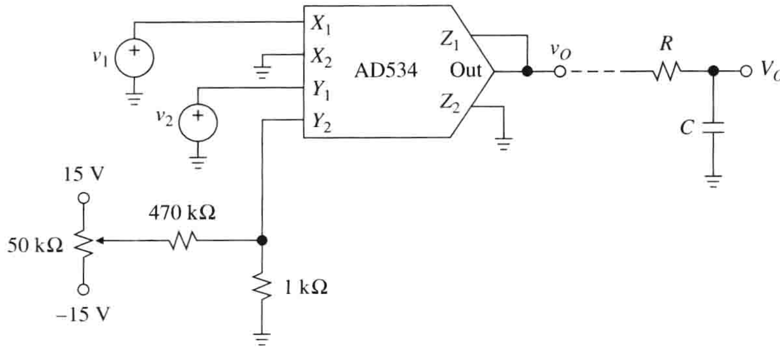


FIGURE 13.10
Basic multiplier connection for $v_O = v_1v_2/10$. If followed by a low-pass filter, it can be used for phase detection.

Figure 13.10 shows the basic connection for signal multiplication, or $v_O = v_1v_2/10$. As such, it forms the basis of amplitude modulation and voltage-controlled amplification. When either input is zero, v_O should also be zero, regardless of the other input. In practice, because of slight component mismatches, a small fraction of the nonzero input will feed through to the output, causing an error. In critical applications such as suppressed-carrier modulation, this error can be minimized by applying an external trim voltage (±30-mV range required) to the X_2 or the Y_2 input.

Of particular interest is the case in which the inputs are ac signals, or $v_1 = V_1\cos(\omega_1 t + \theta_1)$ and $v_2 = V_2\cos(\omega_2 t + \theta_2)$, for then their product is, by a well-known trigonometric identity

$$v_O = \frac{V_1V_2}{20}\{\cos[(\omega_1 - \omega_2)t + (\theta_1 - \theta_2)] + \cos[(\omega_1 + \omega_2)t + (\theta_1 + \theta_2)]\}$$

indicating that v_O consists of two components, with frequencies equal to the sum and the difference of the input frequencies. If the input frequencies are the same and the high-frequency component is suppressed with a low-pass filter, as shown, then we get

$$V_O = \frac{V_1V_2}{20}\cos(\theta_1 - \theta_2) \tag{13.23}$$

In this capacity, the circuit can be used in ac power measurements or as a phase detector in phase-locked-loop circuits.

Figure 13.11 shows how a multiplier can be configured for two other popular functions, namely, *analog division* and *square-root extraction.* In Fig. 13.11*a* we have, by Eq. (13.21), $0 - v_2 = (v_1 - 0)(0 - v_O)/10$, or $v_O = 10(v_2/v_1)$. To maximize the denominator range, return the X_2 input to a trimmable voltage (±3-mV range required).

In Fig. 13.11*b* we have $0 - v_I = (v_O - 0)(0 - v_O)/10$, or $v_O = \sqrt{10v_I}$. The function of the diode is to prevent a latching condition, which could arise in the event of the input inadvertently changing polarity. Additional applications are discussed in the end-of-chapter problems.

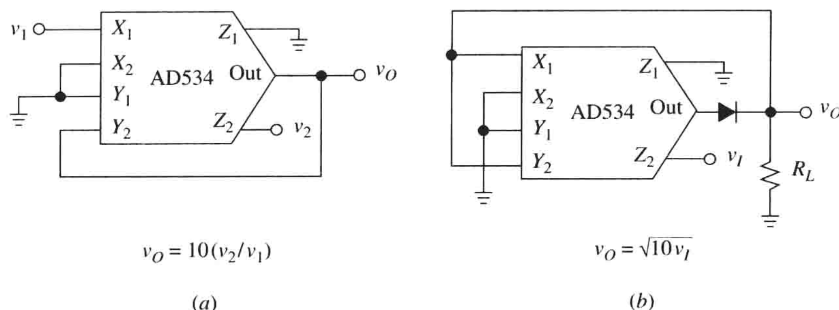


FIGURE 13.11
Analog divider and square rooter.

13.3 OPERATIONAL TRANSCONDUCTANCE AMPLIFIERS

An operational transconductance amplifier (OTA) is a voltage-input, current-output amplifier. Its circuit model is shown in Fig. 13.12*a*. To avoid loading effects both at the input and at the output, an OTA should have $z_d = z_o = \infty$. The ideal OTA, whose circuit symbol is shown in Fig. 13.12*b*, gives $i_O = g_m v_D$, or

$$i_O = g_m(v_P - v_N) \tag{13.24}$$

where g_m is the *unloaded transconductance gain,* in amperes per volt.

In its simplest form, an OTA consists of a differential transistor pair with a current-mirror load.[7] We have encountered this configuration when studying op amp input stages in Chapter 5. In the bipolar example of Fig. 5.1 the OTA consists of the Q_1-Q_2 pair and the Q_3-Q_4 mirror; in the MOS example of Fig. 5.4*a* it consists of the M_1-M_2 pair and the M_3-M_4 mirror.

Besides serving as building blocks for other amplifiers, OTAs find application in their own right. Since it can be realized with just one stage and it operates on the principle of processing currents rather than voltages, the OTA is an inherently fast device.[8] Moreover, g_m can be varied by changing the bias current of the differential transistor pair, making OTAs suited to electronically programmable functions.

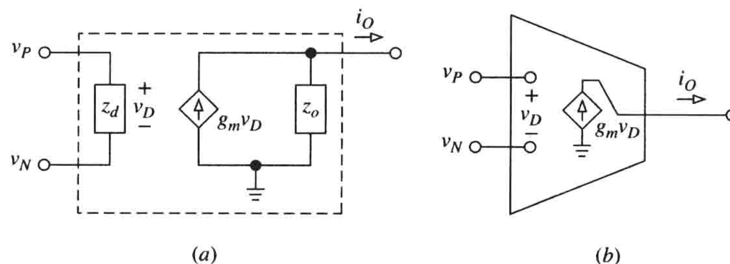


FIGURE 13.12
Operational transconductance amplifier: (*a*) equivalent circuit and (*b*) ideal model.

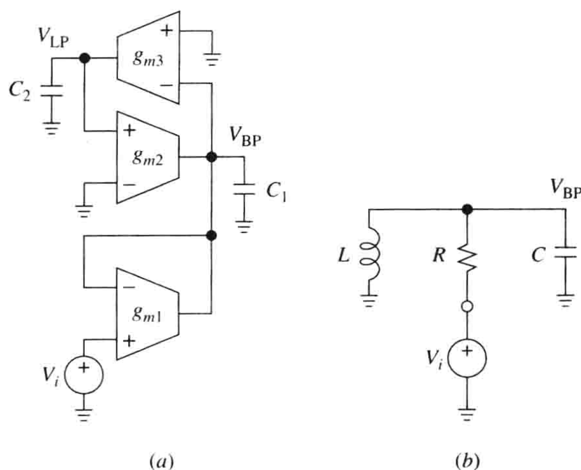


FIGURE 13.13
Second-order g_m-C filter and its *RLC* equivalent.

g_m-C Filters

A popular OTA application is the realization of fully integrated continuous-time filters, where OTAs have emerged as viable alternatives to traditional op amps.[7–9] OTA-based filters are referred to as g_m-C filters because they use OTAs and capacitors, but no resistors and no inductors. A popular g_m-C filter example is shown in Fig. 13.13*a*. Its analysis proceeds as follows.

Denoting the currents out of the three OTAs as I_1, I_2, and I_3, we have, by Eq. (13.24), $I_1 = g_{m1}(V_i - V_{BP})$, $I_2 = g_{m2}V_{LP}$, and $I_3 = -g_{m3}V_{BP}$. By Ohm's law, $V_{LP} = (1/sC_2)I_3$ and $V_{BP} = (1/sC_1)(I_1 + I_2)$. Combining, we get

$$\frac{V_{BP}}{V_i} = \frac{sC_2g_{m1}/(g_{m2}g_{m3})}{s^2C_1C_2/(g_{m2}g_{m3}) + sC_2g_{m1}/(g_{m2}g_{m3}) + 1} \tag{13.25}$$

It is readily seen that this transfer function is the same as that of the *RLC* equivalent of Fig. 13.13*b*, provided $C = C_1$, $R = 1/g_{m1}$, and $L = C_2/(g_{m2}g_{m3})$. Evidently, g_{m1} simulates a resistance, whereas the combination g_{m2}-g_{m3}-C_2 simulates an inductance. Moreover, the circuit provides V_{BP} and V_{LP} simultaneously, a feature not available in its *RLC* counterpart. What is even more important is that we can automatically tune f_0 by varying g_{m2} and g_{m3}, and tune Q by varying g_{m1}.

EXAMPLE 13.2. (*a*) In the filter of Fig. 13.13*a* find g_{m1} and $g_{m2} = g_{m3}$ for $\omega_0 = 10^5$ rad/s and $Q = 5$ with $C_1 = C_2 = 100$ pF. (*b*) What are the values of the simulated resistance and the simulated inductance? (*c*) The sensitivities of the filter?

Solution.

(*a*) By inspection, $\omega_0 = \sqrt{g_{m2}g_{m3}/C_1C_2}$ and $Q = \sqrt{C_1/C_2} \times \sqrt{g_{m2}g_{m3}}/g_{m1}$. Substituting the given data, we get $g_{m2} = g_{m3} = 10\ \mu$A/V and $g_{m1} = 2\ \mu$A/V.
(*b*) $R = 500$ kΩ, and $L = 1$ H.
(*c*) The sensitivity of Q with respect to g_{m1} is -1; all other sensitivities are either $\frac{1}{2}$ or $-\frac{1}{2}$, which are fairly low.

Off-the-Shelf OTAs

Figure 13.14 shows a popular OTA available as an off-the-shelf IC. Its heart is the linearized transconductance multiplier made up of D_1-D_2 and Q_3-Q_4. The remaining blocks, each consisting of a BJT pair and a diode, are high-output-impedance current mirrors of the Wilson type. Denoting the collector current of transistor Q_k as i_k and ignoring base currents, we can describe circuit operation as follows.

The mirror Q_1-D_3-Q_2 accepts the external control current I_C and duplicates it at the emitters of the Q_3-Q_4 pair to give

$$i_3 + i_4 = I_C$$

The mirror Q_5-D_4-Q_6 duplicates i_3 to yield $i_6 = i_3$, and the mirror Q_7-D_5-Q_8 duplicates i_6 to yield $i_8 = i_6$, so $i_8 = i_3$. Likewise, the mirror Q_9-D_6-Q_{10} duplicates i_4 to yield $i_{10} = i_4$. Consequently, KCL gives $i_O = i_{10} - i_8$, or

$$i_O = i_4 - i_3$$

Retracing the reasoning of Section 5.1, we can write

$$i_O = I_C \tanh \frac{v_P - v_N}{2V_T} \tag{13.26}$$

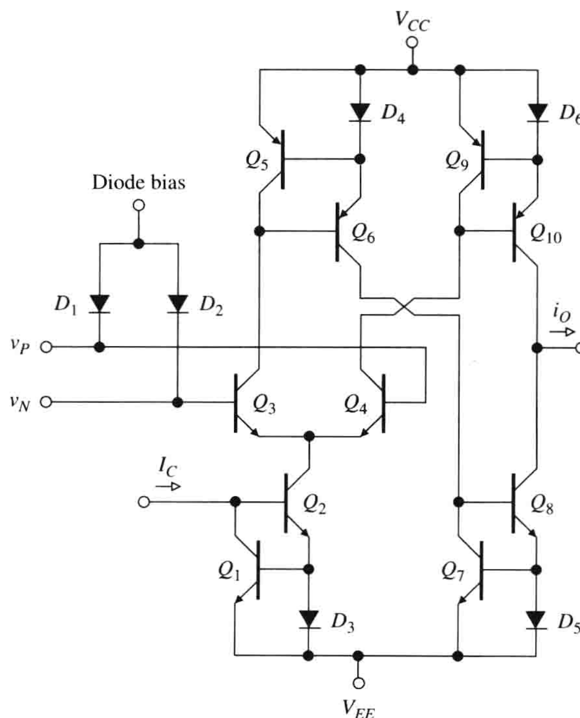


FIGURE 13.14
Bipolar OTA.

where V_T is the thermal voltage ($V_T \cong 26$ mV at room temperature). The plot of i_O versus $(v_P - v_N)$ is the familiar s-shaped curve of Fig. 5.2b. As discussed in connection with Fig. 1.24, the output will generally be distorted.

There are two ways to keep distortion under control. One is to restrict the Q_3-Q_4 pair to *small-signal operation*. To see how, expand as $\tanh x = x - x^3/3 + \ldots$, and observe that to approximate linear operation ($\tanh x \cong x$) we need to impose $|x^3/3| \ll |x|$, or $|x| \ll \sqrt{3}$. Letting $x = (v_P - v_N)/(2V_T)$, small-signal operation requires that we keep $|v_P - v_N| \ll 2\sqrt{3}V_T$ ($\cong 90$ mV). Under this condition we have

$$i_o \cong G_m(v_p - v_n)$$

where the transconductance gain

$$G_m = \frac{I_C}{2V_T} \tag{13.27a}$$

is linearly proportional to the control current I_C; as such, it is linearly programmable by the user. As we know, BJTs handle current ranges in excess of five decades, or more than 100 dB, making bipolar OTAs particularly suited to a variety of programmable circuits such as programmable amplifiers, oscillators, and filters, especially in wide-dynamic range applications like audio and electronic music.

The small-signal constraint is generally a drawback, so a viable alternative is to *predistort* the input drive $(v_P - v_N)$ according to the *inverse* of the tanh function. A circuit approximating this function is the diode network of Fig. 13.15a, consisting of: (a) the voltage dividers R_1-R_2 and R_3-R_4 to suitably scale the external inputs (in Fig 13.15a only one input, v_I, is used, and the resistance values are designed to accommodate input peak amplitudes of ± 5 V); (b) the nonlinear diode pair D_1-D_2;

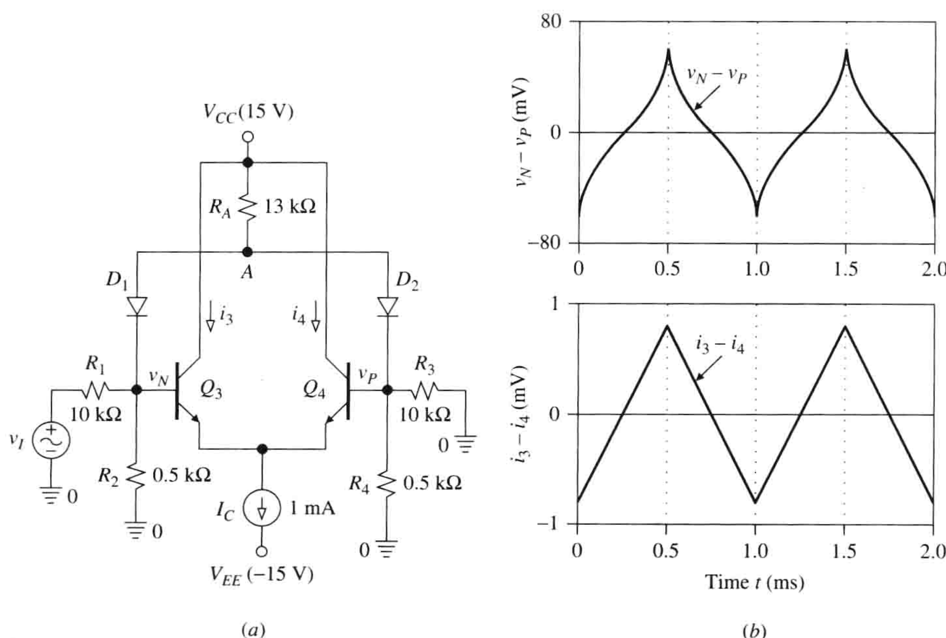


FIGURE 13.15
Using PSpice to visualize the linearizing effect of the input diode network.

(*c*) the diode-biasing resistance R_A. Figure 13.15*b* shows how the diode network converts a triangular input v_I of ±5-V peak values to a suitably distorted drive $(v_N - v_P)$ of ±60-mV peak values to yield a virtually undistorted triangular output $(i_3 - i_4)$, which the Wilson mirrors then convert to the output current i_O. (Note the similarity of Fig. 13.15*b* to Fig. 1.26*b*, except that the predistortion of Fig. 1.26 is achieved via negative feedback, whereas that of Fig. 13.15 is achieved by means of the nonlinear diode network.) In light of the above, we express the transfer characteristic of the circuit of Fig. 13.15*a* as

$$i_3 - i_4 = G_m v_I$$

where the transconductance gain is found empirically to be, for the component values shown,

$$G_m \cong \frac{1}{12.2\ \mathrm{k}\Omega} \tag{13.27b}$$

Applications with Off-the-Shelf OTAs

Though OTA data sheets propose quite a variety of useful applications, we shall examine a few representative ones, namely, voltage-controlled amplifiers, filters, and oscillators (VCAs, VCFs, and VCOs).

Figure 13.16 shows a basic VCA (note the alternative symbol for the OTA). Here OA_1 and Q_1 form a *V-I* converter to provide $I_C = V_C/R$, where we are assuming the base current of Q_1 to be negligible. OA_2 converts i_O to a voltage v_O, and since i_O is proportional to the product $I_C \times v_I$, the final result is $v_O = Av_I$,

$$A = kV_C \tag{13.28}$$

where k is a suitable proportionality constant, in V^{-1}. The 1-kΩ pot is used for offset nulling, and the 25-kΩ pot for the calibration of k. By Eq. (13.27*b*), adjusting the 25-kΩ pot near 12.2 kΩ yields $k = 1/(10\ \mathrm{V})$, indicating that varying V_C from 0 to

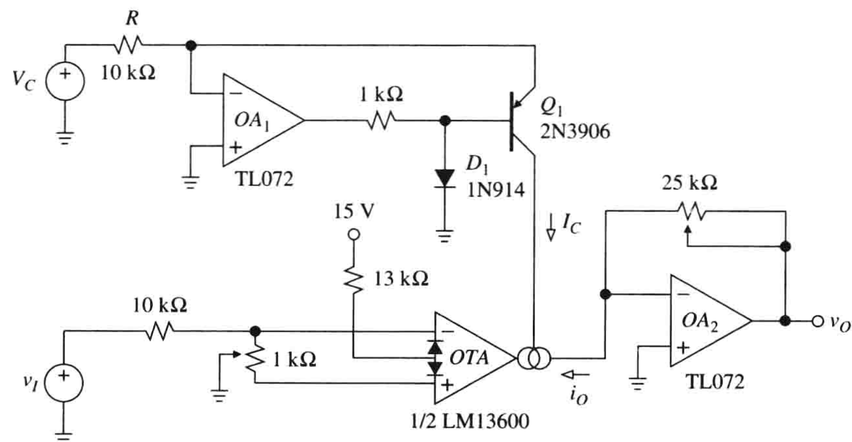


FIGURE 13.16
Voltage-controlled amplifier with linear control.

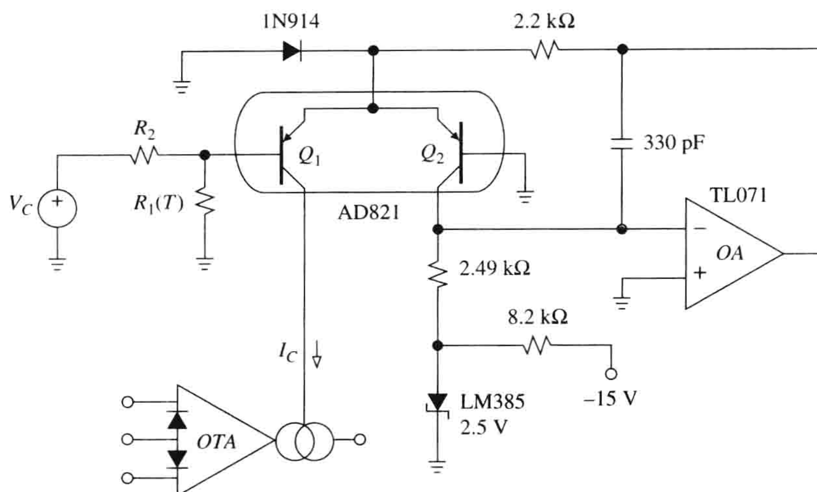


FIGURE 13.17
OTA with exponential control.

10 V will change A from 0 to 1 V/V. The circuit is calibrated as follows: (a) with $v_I = 0$, sweep V_C from 0 to 10 V and adjust the 1-kΩ pot for the minimum deviation of v_O from 0 V; (b) with $V_C = 10$ V, adjust the 25-kΩ pot so that $v_O = 10$ V for $v_I = 10$ V.

The circuit of Fig. 13.16 provides linear gain control. Audio applications often call for exponential control, or

$$A = A_0 b^{kV_C} \tag{13.29}$$

where b is usually either 10 or 2, k is a proportionality constant in decades or octaves per volt, and A_0 is the gain for $V_C = 0$. Exponential control is readily achieved by generating I_C with an antilog converter, as shown in Fig. 13.17. Since I_C must be sourced to the OTA, the BJTs must be of the *pnp* type. With $V_C = 0$, the circuit gives $I_C = 1$ mA; increasing V_C decreases I_C exponentially with a sensitivity of k dec/V or k oct/V, with k being set by R_2.

OTA-based VCFs and VCOs rely on the integration of the OTA's output current using a capacitor. The example of Fig. 13.18 also uses an op amp to provide low output

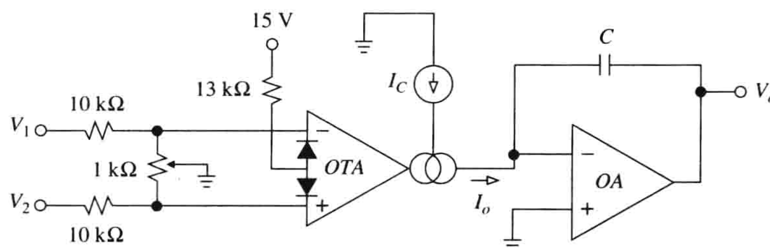


FIGURE 13.18
Current-controlled integrator.

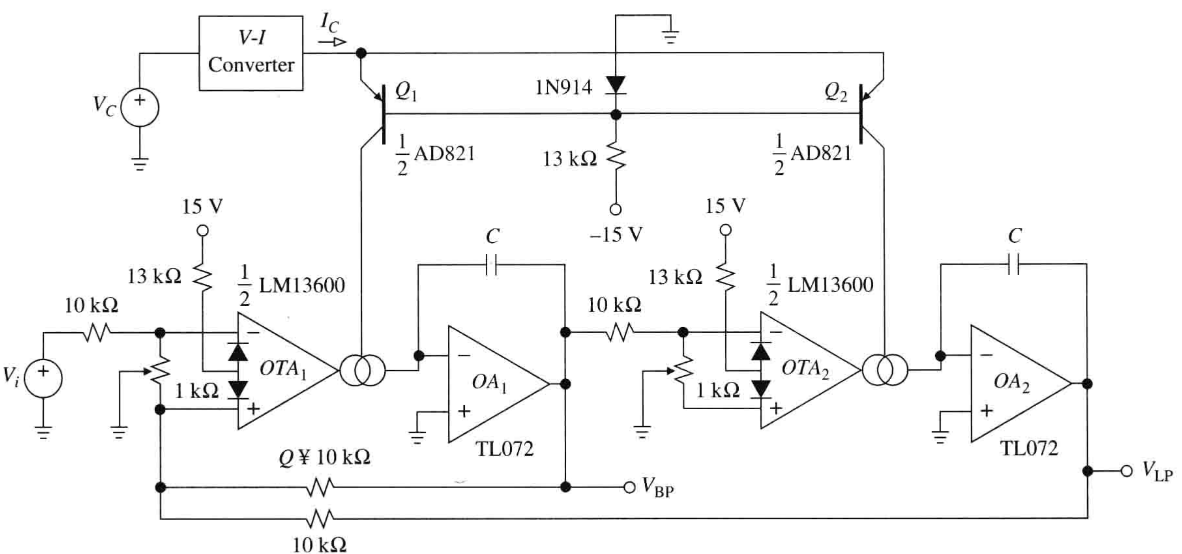


FIGURE 13.19
Voltage-controlled state-variable filter.

impedance. Writing $V_o = (-1/sC)I_o = (-1/sC) \times [I_C/(12.2\ \text{V})] \times (V_2 - V_1)$, and letting $s = j2\pi f$, we obtain

$$V_o = \frac{1}{jf/f_0}(V_1 - V_2) \qquad f_0 = \frac{I_C}{2\pi(12.2\ \text{V})C} \tag{13.30}$$

The circuit integrates the difference $V_1 - V_2$ with a programmable unity-gain frequency f_0. For instance, varying I_C from 1 μA to 1 mA with $C = 652$ pF will sweep f_0 over the entire audio range, from 20 Hz to 20 kHz. I_C can be generated either with a linear *V-I* converter, as in Fig. 13.16, or with an exponential converter, as in Fig. 13.17.

The circuit of Fig. 13.19 uses two OTA-based integrators to implement a state-variable topology of the type of Fig. 4.37. The output current of the *V-I* converter, which can be controlled either linearly or exponentially, is split between the two OTAs by the suitably biased AD821 matched BJT pair.

Applying Eq. (13.30), we get $V_{BP} = (V_i - V_{BP}/Q - V_{LP})/(jf/f_0)$ and $V_{LP} = V_{BP}/(jf/f_0)$. Combining, we obtain

$$\frac{V_{BP}}{V_i} = QH_{BP} \qquad \frac{V_{LP}}{V_i} = H_{LP} \tag{13.31}$$

where H_{BP} and H_{LP} are the standard second-order band-pass and low-pass functions defined in Section 3.4, and

$$f_0 = \frac{I_C}{4\pi(12.2\ \text{V})C}$$

If a resonance gain of unity is desired, then increase the 10-kΩ input resistance by a factor of Q. To reduce Q-enhancement effects, follow the directions of Section 6.5 and use a small phase-lead capacitor in parallel with the 10-kΩ interstage resistance.

In the circuit of Fig. 13.20 the OTA is used to source/sink a current of value I_C, and thus charge/discharge C at a programmable rate. The resulting triangle waveform alternates between 5 V and 10 V, the thresholds of the high-input-impedance CMOS

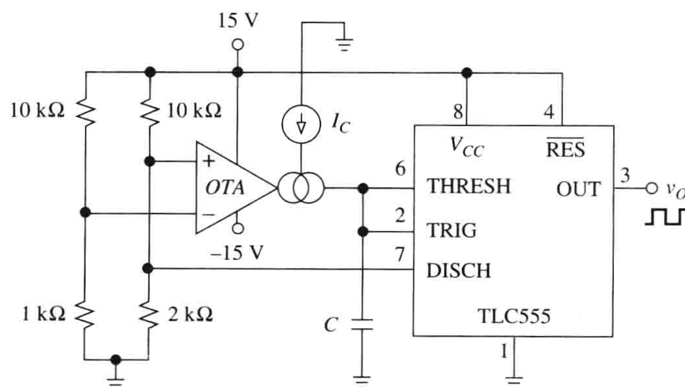


FIGURE 13.20
Current-controlled relaxation oscillator.

timer. The frequency of oscillation is (see Problem 13.16),

$$f_0 = \frac{I_C}{10C} \tag{13.32}$$

As usual, I_C can be controlled either linearly or exponentially. If the triangular wave is used, a buffer amplifier may be required.

13.4 PHASE-LOCKED LOOPS

A *phase-locked loop* (PLL) is a frequency-selective circuit designed to synchronize with an incoming signal and maintain synchronization in spite of noise or variations in the incoming signal frequency. As depicted in Fig. 13.21, the basic PLL system comprises a *phase detector,* a *loop filter,* and a *voltage-controlled oscillator* (VCO).

The phase detector compares the phase θ_I of the incoming signal v_I against the phase θ_O of the VCO output v_O, and develops a voltage v_D proportional to the difference $\theta_I - \theta_O$. This voltage is sent through a low-pass filter to suppress high-frequency ripple and noise, and the result, called the *error voltage* v_E, is applied to the control input of the VCO to adjust its frequency ω_O.

The VCO is designed so that with $v_E = 0$ it is oscillating at some initial frequency ω_0 called the *free-running frequency,* so its characteristic is

$$\omega_O(t) = \omega_0 + K_o v_E(t) \tag{13.33}$$

where K_o is the *sensitivity* of the VCO, in radians-per-second per volt. If a periodic input is applied to the PLL with frequency ω_I sufficiently close to the free-running frequency ω_0, an error voltage v_E will develop, which will adjust ω_O until v_O becomes synchronized with v_I, that is, until *for every input cycle there is one, and only one, VCO cycle.* At this point the PLL is said to be *locked* on the incoming signal, and it gives $\omega_O = \omega_I$ exactly.

Should ω_I change, the phase shift between v_O and v_I will start to increase, changing v_D and, hence, the control voltage v_E. This change in v_E is designed to adjust the VCO until ω_O is brought back to the same value as ω_I. This self-adjusting ability by the feedback loop allows the PLL, once locked, to track input frequency changes. Since a change in ω_I is ultimately reflected by a change in v_E, we use v_E as the output of the PLL whenever we wish to detect changes in ω_I, as in FM and FSK demodulation.

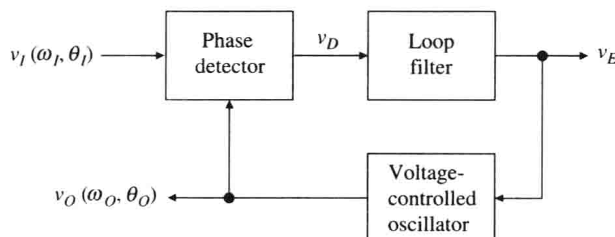


FIGURE 13.21
Basic phase-locked loop.

A PLL can be designed to lock on the incoming signal in spite of noise that might afflict such a signal. A noisy input will generally cause the phase-detector output v_D to jitter around some average value. However, if the filter cutoff frequency is low enough to suppress this jitter, v_E will emerge as a clean signal, in turn resulting in a stable VCO frequency and phase. We thus use ω_O as the output of the PLL whenever we wish to recover a signal buried in noise, and also in frequency-related applications such as frequency synthesis and synchronization.

Lock and Capture

To develop a concrete understanding of PLL operation, consider the case of phase detection being accomplished with a balanced mixer of the type discussed in Section 13.2. As we know, the mixer output contains the sum and difference frequencies $\omega_I \pm \omega_O$. When the loop is locked, the sum is twice ω_I and the difference is zero, or dc. The low-pass filter suppresses the sum but passes the dc component, which thus keeps the loop in lock.

If the loop is not locked, and if the difference frequency falls above the cutoff frequency of the filter, it will be suppressed along with the sum frequency, leaving the loop unlocked and oscillating at its free-running frequency. However, if ω_O is sufficiently close to ω_I to make the difference frequency approach the filter band edge, part of this component is passed, tending to drive ω_O toward ω_I. As the difference $\omega_O - \omega_I$ is reduced, more error signal is transmitted to the VCO, resulting in a constructive effect that ultimately brings the PLL in lock.

The *capture range* is the frequency range $\pm\Delta\omega_C$, centered about ω_0, over which the loop can acquire lock. This range is affected by the filter characteristics, and gives an indication of how close ω_I must be to ω_0 to acquire lock. The *lock range* is the frequency range $\pm\Delta\omega_L$, also centered about ω_0, over which the loop can track the input once lock has been established. The lock range is affected by the operating range of the phase detector and the VCO. The capture process is a complex phenomenon, and the capture range is never greater than the lock range.

The time it takes for a PLL to capture the incoming signal is called the *capture time* or *pull-in time.* This time depends on the initial frequency and phase differences between v_I and v_O, as well as the filter and other loop characteristics. In general, it can be said that reducing the filter bandwidth has the following effects: (a) it slows down the capture process, (b) it increases the pull-in time, (c) it decreases the capture range, and (d) it increases the interference-rejection capabilities of the loop.

The PLL in the Locked Condition

When in the locked condition, a PLL can be modeled[10–12] as in Fig. 13.22. This diagram is similar to that of Fig. 13.21, except that we are now working with Laplace transforms of *signal changes* (symbolized by lowercase letters with lowercase subscripts) about some operating point, and *operations* on these changes, both of which are generally functions of the complex frequency s. The phase detector develops the voltage change

$$v_d(s) = K_d\theta_d(s) \tag{13.34a}$$

$$\theta_d(s) = \theta_i(s) - \theta_o(s) \tag{13.34b}$$

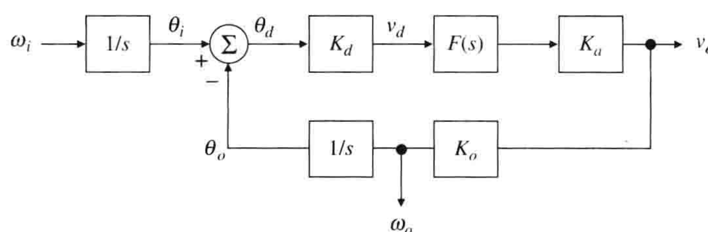


FIGURE 13.22
Block diagram of the basic PLL system in the locked condition.

where K_d is the phase-detector *sensitivity,* in volts per radian. This voltage is sent through the loop filter, whose transfer function is denoted as $F(s)$, and possibly an amplifer with gain K_a, in volts per volt, to produce the error-voltage variation $v_e(s)$. This, in turn, is converted to the frequency variation $\omega_o(s) = K_o v_e(s)$, by Eq. (13.33).

Since the phase detector processes phase, we need a means for converting from frequency to phase. Considering that frequency represents the rate of change of phase with time, or $\omega = d\theta(t)/dt$, we have

$$\theta(t) = \theta(0) + \int_0^t \omega(\xi)\, d\xi \tag{13.35}$$

indicating that frequency-to-phase conversion is inherently an operation of integration. Exploiting the well-known Laplace transform property that integration in the time domain corresponds to division by s in the frequency domain, we use the $1/s$ blocks shown.

If we were to open the loop at the inverting input of the phase comparator, the overall gain experienced by $\theta_i(s)$ in going around the path and emerging as $\theta_o(s)$ is $K_d \times F(s) \times K_a \times K_o \times 1/s$, or

$$T(s) = K_v \frac{F(s)}{s} \tag{13.36}$$

$$K_v = K_d K_a K_o \tag{13.37}$$

where $T(s)$ is the *open-loop gain,* in radians per radian, and K_v is called the *gain factor,* in s^{-1}. With the loop closed, we readily find

$$H(s) = \frac{\theta_o(s)}{\theta_i(s)} = \frac{T(s)}{1 + T(s)} = \frac{K_v F(s)}{s + K_v F(s)} \tag{13.38}$$

Other transfer functions may be of interest, depending on what we consider as input and output. For instance, substituting $\theta_i(s) = \omega_i(s)/s$ and $\theta_o = (K_o/s)v_e(s)$, we readily get

$$\frac{v_e(s)}{\omega_i(s)} = \frac{1}{K_o} H(s) \tag{13.39}$$

which allows us to find the voltage change $v_e(s)$ in response to an input frequency change $\omega_i(s)$, as in FM and FSK demodulation.

Comparing Fig. 13.22 with Fig. 1.22, we observe that a PLL is a negative-feedback system with $x_i = \theta_i$, $x_f = \theta_o$, and $a\beta = T = K_v F(s)/s$, indicating that the open-loop gain T plays also the role of the *loop gain* of the system. Further, even though we are more attuned to frequency, we must recognize that the natural input of a PLL is phase. Since $T \to \infty$ as $s \to 0$, the PLL will force θ_o to track θ_i, just as an op amp voltage follower forces v_o to track v_i. In this respect, it pays to view a PLL as a *phase follower.* The fact that it also forces ω_O to track ω_I is a consequence of this phase-follower action, along with the phase-frequency relationship $\omega = d\theta/dt$.

As seen in Chapter 8, the loop gain T affects both the dynamics and the stability of the PLL. In turn, $T(s)$ is strongly influenced by $F(s)$. We make the following observations: (*a*) The number of poles of $H(s)$ defines the *order* of the loop; (*b*) the number of $1/s$ terms (or integrations) present within the loop defines the *type* of loop. Because of the $1/s$ function associated with the VCO, a PLL is at least Type I, and its order equals the order of the filter plus 1.

First-Order Loop

Consider the instructive case in which there is no loop filter, or $F(s) = 1$. The result is a *first-order loop,* and the above equations simplify, after the substitution $s \to j\omega$, as

$$T(j\omega) = \frac{1}{j\omega/K_v} \tag{13.40}$$

$$\frac{v_e(j\omega)}{\omega_i(j\omega)} = \frac{1/K_o}{1 + j\omega/K_v} \tag{13.41}$$

Equation (13.40) indicates a *Type I* loop with crossover frequency $\omega_x = K_v$ and phase margin $\phi_m = 90°$. Equation (13.41) indicates that the loop inherently provides a first-order low-pass response with a dc gain of $1/K_o$ V/(rad/s) and a cutoff frequency of K_v rad/s.

If $\omega_i(t)$ is a step change, the resulting change $v_e(t)$ will be an exponential transient governed by the time constant $\tau = 1/K_v$. If $\omega_i(t)$ is varied sinusoidally with a modulating frequency ω_m, $v_e(t)$ will also vary sinusoidally with the same frequency ω_m; its amplitude is $|v_e| = (1/K_o)|\omega_i|$ at low frequencies, and rolls off with ω_m at the rate of -1 dec/dec past K_v.

EXAMPLE 13.3. A first-order PLL with $K_v = 10^4$ s^{-1} uses a VCO with a free-running frequency of 10 kHz and a sensitivity of 5 kHz/V. (*a*) What is the control voltage needed to lock the PLL on a 20-kHz input signal? On a 5-kHz input signal? (*b*) Find the response $v_e(t)$ if the input frequency is changed stepwise as $f_I = [10 + u(t)]$ kHz, where $u(t) = 0$ for $t < 0$ and $u(t) = 1$ for $t > 0$. (*c*) Repeat if the input frequency is modulated as $f_I = 10[1 + 0.1\cos(2\pi f_m t)]$ kHz, $f_m = 2.5$ kHz.

Solution.

(*a*) By Eq. (13.33), $v_E = (\omega_O - \omega_0)/K_o$, where $\omega_0 = 2\pi 10^4$ rad/s and $K_o = 2\pi \times 5 \times 10^3 = \pi 10^4$ (rad/s)/V. For $\omega_O = 2\pi \times 20 \times 10^3$ rad/s we get $v_E = 2$ V, and for $\omega_O = 2\pi \times 5 \times 10^3$ rad/s we get $v_E = -1$ V.

(*b*) The response to a step increase $\omega_i(t) = 2\pi u(t)$ krad/s is an exponential transient with amplitude $|\omega_i(t)|/K_o = 2\pi 10^3/10^4\pi = 0.2$ V, and time constant $1/K_v = 1/10^4 = 100\ \mu$s, so

$$v_e(t) = 0.2[1 - e^{-t/(100\ \mu s)}]u(t)\ \text{V}$$

(*c*) Now $\omega_i(t) = 2\pi \times 10^4 \times 0.1(\cos 2\pi f_m t) = 2\pi 10^3 \cos(2\pi 2500t)$ rad/s. Calculating Eq. (13.41) at $j\omega = j\omega_m = j2\pi 2500$ rad/s gives

$$\frac{v_e(j\omega_m)}{\omega_i(j\omega_m)} = \frac{1/10^4\pi}{1 + j2\pi 2500/10^4} = \frac{0.5370}{10^4\pi} \underline{/-57.52^\circ}\ \text{V/(rad/s)}$$

Letting $\omega_i(j\omega_m) = 2\pi 10^3 \underline{/0^\circ}$ rad/s gives $v_e(j\omega_m) = 0.1074 \underline{/-57.52^\circ}$ V, so

$$v_e(t) = 0.1074\cos(2\pi 2500t - 57.52^\circ)\ \text{V}$$

The absence of a loop filter drastically limits the selectivity and noise-suppression capabilities of a PLL, so first-order loops are seldom used in practice.

Second-Order Loops

Most PLLs utilize a one-pole low-pass filter and are thus *second-order loops.* Such a filter provides a flywheel-like function that allows the VCO to smooth over noise and jumps in the input frequency. As seen in Chapter 8, the presence of a second pole within the loop erodes the phase margin, so care must be exercised to avoid instability. Second-order loops are also stabilized by introducing a filter zero to counterbalance the phase lag due to the filter pole.

A popular loop filter is shown in Fig. 13.23*a*. Called a *passive lag-lead filter,* it provides the transfer function

$$F(s) = \frac{1 + s/\omega_z}{1 + s/\omega_p} \tag{13.42}$$

where $\omega_z = 1/(R_2C)$ and $\omega_p = 1/(R_1 + R_2)C$. By Eq. (13.36), the loop gain is now

$$T(j\omega) = \frac{1 + j\omega/\omega_z}{(j\omega/K_v)(1 + j\omega/\omega_p)} \tag{13.43}$$

indicating a *second-order, Type I loop.* This gain is plotted in Fig. 13.23*b* for the case in which ω_z is positioned at the geometric mean of ω_p and K_v, or $\omega_z = \sqrt{\omega_p K_v}$. The crossover frequency is then ω_z itself, and the phase margin is 45°. Also shown for comparison is the loop gain of the first-order loop, or $F(s) = 1$.

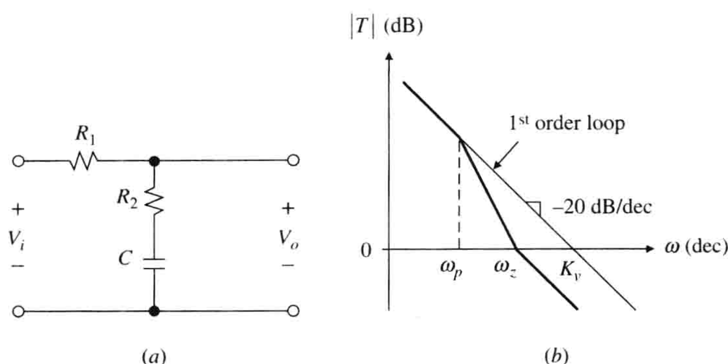


FIGURE 13.23
Passive lag-lead filter, and magnitude plot of the loop gain T.

EXAMPLE 13.4. (*a*) Given a PLL system with $K_v = 10^4\ \text{s}^{-1}$, specify a passive lag-lead filter for a crossover frequency $\omega_x = 10^3$ rad/s and a phase margin $\phi_m = 45°$. (*b*) What are the actual values of ω_x and ϕ_m?

Solution.

(*a*) For $\phi_m \cong 45°$ we want $\omega_z = \omega_x = 10^3$ rad/s, so $\omega_p = \omega_z^2/K_v = 10^6/10^4 = 100$ rad/s. Let $C = 0.1\ \mu\text{F}$. Then, $R_2 = 1/(\omega_z C) = 10\ \text{k}\Omega$, and $R_1 = 1/(\omega_p C) - R_2 = 90\ \text{k}\Omega$ (use 91 kΩ).

(*b*) Using Eq. (13.43), along with the trial-and-error technique of Example 8.1, we find the actual values $\omega_x = 1.27$ krad/s, and $\phi_m = 180° + \measuredangle T(j1.27 \times 10^3) = 56°$.

Another popular loop filter is the *active PI filter* of Fig. 13.24*a*, so called because its output is proportional to the input as well as to the integral of the input. The inverting stage OA_2 can be omitted by interchanging the phase-detector inputs, if needed. Assuming ideal op amps, the filter gives

$$F(s) = \frac{1 + s/\omega_z}{s/\omega_p} \tag{13.44}$$

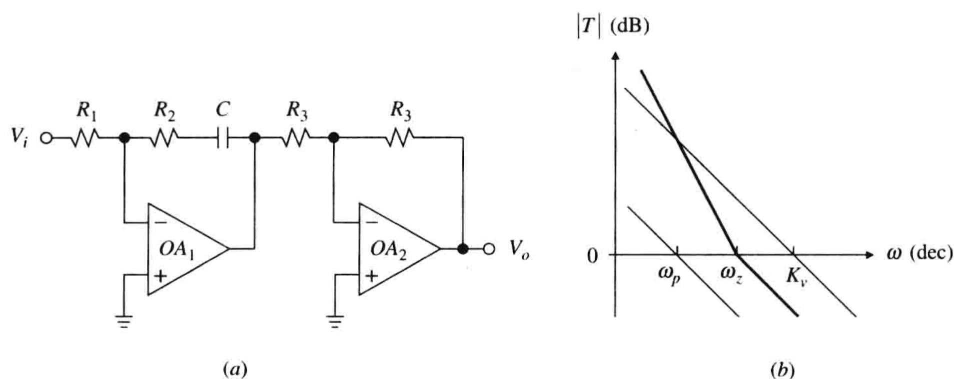


FIGURE 13.24
Active PI filter, and magnitude plot of the loop gain T.

where $\omega_z = 1/R_2C$ and $\omega_p = 1/R_1C$. The corresponding loop gain is

$$T(j\omega) = \frac{1 + j\omega/\omega_z}{(j\omega/K_v)(j\omega/\omega_p)} \tag{13.45}$$

indicating a *second-order, Type II loop*. As shown in Fig. 13.24*b*, the slope is -40 dB/dec below ω_z, and -20 dB/dec above ω_z. Imposing again $\omega_z = \sqrt{\omega_p K_v}$ gives $\omega_x \cong \omega_z$ and $\phi_m \cong 45°$.

Compared to the passive filter, whose dc gain is $F(0) = 1$, the active filter gives $F(0) = \infty$, indicating that the phase error θ_D needed to sustain the control voltage v_E approaches zero when the PI filter is used. In practice, $F(0)$ is limited by the finite dc gain of OA_1; even so, θ_D will still be vanishingly small, implying $\theta_O \cong \theta_I$, or phase coherence at the detector input. The use of the active filter also avoids possible loading effects at the output.

Damping Characteristics

Additional insight is gained by substituting Eqs. (13.42) and (13.44) into Eq. (13.38), and then expressing the latter in the standard form of Eq. (3.40). The results (see Problem 13.20) are

$$H(s) = \frac{(2\zeta - \omega_n/K_v)(s/\omega_n) + 1}{(s/\omega_n)^2 + 2\zeta(s/\omega_n) + 1} \tag{13.46a}$$

$$\omega_n = \sqrt{\omega_p K_v} \qquad \zeta = \frac{\omega_n}{2\omega_z}\left(1 + \frac{\omega_z}{K_v}\right) \tag{13.46b}$$

for the passive lag-lead filter, and

$$H(s) = \frac{2\zeta(s/\omega_n) + 1}{(s/\omega_n)^2 + 2\zeta(s/\omega_n) + 1} \tag{13.47a}$$

$$\omega_n = \sqrt{\omega_p K_v} \qquad \zeta = \frac{\omega_n}{2\omega_z} \tag{13.47b}$$

for the active PI filter. As we know, ω_n is the *undamped natural frequency,* and ζ is the *damping ratio.* If $\omega_n \ll K_v$, as is predominantly the case, Eq. (13.46) reduces to Eq. (13.47) and the PLL with the passive lag-lead filter is said to be a *high-gain loop.* We observe that $H(s)$ is in both cases a combination of the band-pass response H_{BP} and the low-pass response H_{LP}. At low frequencies $H \to H_{LP}$, but at high frequencies $H \to H_{BP}$.

Recall that for $\zeta < 1$ the step response exhibits overshoot. To keep the latter within reason, it is customary to design for $0.5 \le \zeta \le 1$. Under this condition, the time constant governing the loop response to small phase or frequency changes is roughly[7]

$$\tau \cong \frac{1}{\omega_n} \tag{13.48}$$

and the loop bandwidth, obtained by imposing $|H(j\omega)| = 1/\sqrt{2}$, is[13]

$$\omega_{-3\text{ dB}} = \omega_n[1 \pm 2\zeta^2 + \sqrt{1 + (1 \pm 2\zeta^2)^2}]^{1/2} \tag{13.49}$$

where the plus (minus) sign holds for high-gain (low-gain) loops.

EXAMPLE 13.5. (*a*) Find ζ, τ, and $\omega_{-3\text{ dB}}$ for the PLL of Example 13.4. (*b*) Find the response $v_e(t)$ to small input changes of the type $\omega_i = |\omega_i|u(t)$ and $\omega_i = |\omega_i|\cos(\omega_m t)$, $\omega_m = 1$ krad/s.

Solution.

(*a*) By Eq. (13.46*b*), $\omega_n = \sqrt{10^2 \times 10^4} = 1$ krad/s and $\zeta = [10^3/(2 \times 10^3)](1 + 10^3/10^4) = 0.55$. Using Eq. (13.49) for the high-gain loop case gives $\omega_{-3\text{ dB}} \cong 1.9\omega_n = 1.9$ krad/s. By Eq. (13.48), $\tau \cong 1/10^3 = 1$ ms.

(*b*) Substituting the above data into Eq. (13.46*a*) gives

$$H(s) = \frac{s/10^3 + 1}{(s/10^3)^2 + 1.1(s/10^3) + 1}$$

This function has a complex pole pair at $s = -550 \pm j835$ complex Np/s, indicating a step response of the type

$$v_e(t) = \frac{|\omega_i|}{K_o}[1 - Ae^{-550t}\cos(835t + \phi)]$$

with A and ϕ suitable constants. Calculating $H(s)$ at $s = j\omega_m$ as in Example 13.3, we find the ac response as

$$v_e(t) = \frac{|\omega_i|}{K_o}1.286\cos(10^3 t - 45^\circ)$$

Filter Design Criteria

In general, ω_n is chosen high enough to ensure satisfactory dynamics, yet low enough to provide sufficient flywheel action for smoothing over undesired frequency jumps or noise. A typical design process proceeds as follows: (*a*) first, choose ω_n to achieve either the desired $\omega_{-3\text{ dB}}$ or the desired τ, depending on the application; (*b*) next, using Eq. (13.46*b*) or (13.47*b*), specify ω_p for the chosen ω_n; (*c*) finally, specify ω_z for the desired ζ.

We observe that because of the filter zero, a second-order PLL acts as a first-order loop at high frequencies, indicating a reduced ability to suppress ripple and noise. This drawback can be overcome by adding a capacitance $C_2 \ll C$ in parallel with R_2 in either of the above filters. This creates an additional high-frequency pole and turns the loop into a *third-order loop*. To avoid perturbing the existing values of ω_x and ϕ_m significantly, this pole is positioned about a decade above ω_x by imposing $1/(R_2C_2) \cong 10\omega_x$.

EXAMPLE 13.6. Redesign the filter of Example 13.4 for $\omega_{-3\text{ dB}} = 1$ krad/s and $\zeta = 1/\sqrt{2}$. What are the new values of τ and ϕ_m? What value of C_2 would yield a third-order loop without reducing ϕ_m too much?

Solution. With $\zeta = 1/\sqrt{2}$ we get $\omega_n \cong \omega_{-3\text{ dB}}/2 = 10^3/2.0 = 500$ rad/s, so $\tau \cong 2$ ms. Equation (13.46*b*) gives $\omega_p = 25$ rad/s and $\omega_z = 366$ rad/s, which can be realized with $C = 1\ \mu$F, $R_1 = 39$ kΩ, and $R_2 = 2.7$ kΩ. Proceeding as in Example 13.4, we find $\omega_x \cong 757$ rad/s, and $\phi_m \cong 66^\circ$. Use $C_2 \cong C/10 = 0.1\ \mu$F.

13.5 MONOLITHIC PLLs

Monolithic PLLs are available in various technologies and in a wide range of performance specifications.[12] In the following we discuss the popular 4046 CMOS PLL as a representative example.

The 74HC(T)4046A CMOS PLL

Originally developed by RCA, the 4046 family of CMOS PLLs has gone through a series of improvements, and presently includes the 74HC(T)4046A, the 74HC(T)7046A, and the 74HC(T)9046A.[13] We select the 4046A version, shown in simplified form in Fig. 13.25, because it includes the three most common phase detector types, known as *Type I* (PC_1), *Type II* (PC_2), and *Type III* (PC_3) *phase comparators*. Since the circuit is powered from a single supply (typically $V_{SS} = 0$ V and $V_{DD} = 5$ V), all analog signals are referenced to $V_{DD}/2$, or 2.5 V.

The VCO

The VCO, whose details[7,13] are omitted for brevity, is a current-controlled multivibrator operating on a principle similar to that of the emitter-coupled VCO of Fig. 10.30. The current for the capacitor is obtained from the control voltage v_E via a *V-I* converter whose sensitivity is set by R_1 and whose output is offset by R_2. The VCO characteristic is of the type

$$f_O = \frac{k_1}{R_1 C} v_E + \frac{k_2}{R_2 C} \tag{13.50}$$

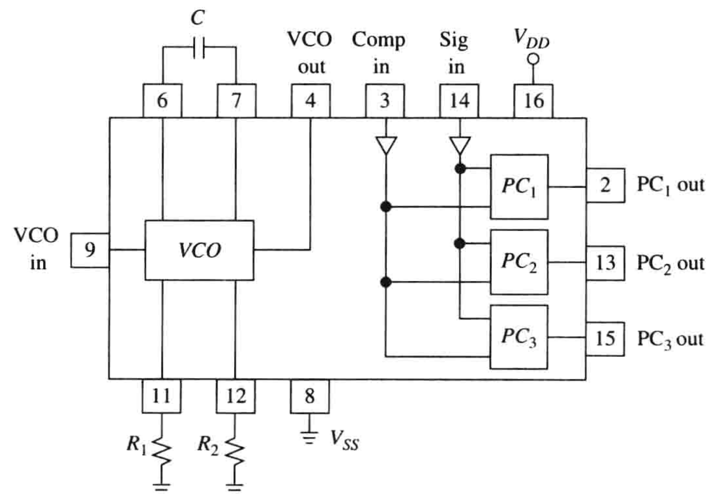


FIGURE 13.25
Simplified block diagram of the 4046A CMOS PLL.

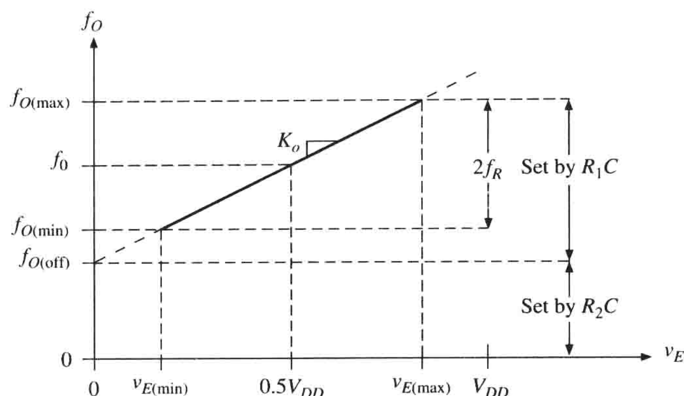


FIGURE 13.26
VCO characteristic and terminology.

where k_1 and k_2 are suitable circuit constants. As shown in Fig. 13.26, the value of f_O corresponding to $v_E = V_{DD}/2$ is called the *center frequency* f_0. It is apparent that if R_2 is omitted ($R_2 = \infty$), the *frequency offset* $f_{O(\text{off})} = k_2/(R_2C)$ becomes zero. The maximum VCO frequency of CMOS PLLs is typically on the order of 10 MHz.

The VCO characteristic of Eq. (13.50) holds only as long as v_E is confined within the range $v_{E(\min)} \le v_E \le v_{E(\max)}$. For a 4046A PLL with $V_{DD} = 5$ V, this range is typically[13] $1.1\text{ V} \le v_E \le 3.9$ V. The frequency range corresponding to the permissible range of v_E is called the *VCO frequency range* $2f_R$. Outside this range the VCO characteristic depends on the particular 4046 version, and it can be found in the data sheets.

The VCO sensitivity is $K_o = 2f_R/[v_{E(\max)} - v_{E(\min)}]$. In FM applications it is usually required that the *V-F* characteristic of the VCO be highly linear in order to minimize distortion. However, in such applications as frequency synchronization, synthesis, and reconstruction the linearity requirements are less stringent.

The Type I Phase Comparator

The Type I phase comparator, depicted in Fig. 13.27*a*, is an exlusive-OR (XOR) gate. This gate outputs $v_D = V_{DD} = 5$ V whenever its input levels disagree with each other, and $v_D = V_{SS} = 0$ whenever they agree. This is exemplified in the timing diagram of Fig. 13.28, where the waveforms have been plotted as a function of $\omega_I t$. It is apparent that if we average out $v_D(t)$ by means of a low-pass filter, the result is $V_D = DV_{DD}$, where D is the duty cycle of v_D. D is minimized when the inputs are in phase with each other, and maximized when they are in antiphase. If both input waveforms have 50% duty cycles, as shown, then $0 \le D \le 1$. Consequently, PC_1 will exhibit the characteristic of Fig. 13.27*b*, and $K_d = V_{DD}/\pi = 5/\pi = 1.59$ V/rad.

An alternative implementation of the Type I comparator, especially in bipolar PLLs designed to work with low-amplitude inputs, is a four-quadrant multiplier, as discussed in Section 13.2. Also called a *balanced modulator,* the multiplier is implemented with a scale factor high enough to ensure that v_I will typically

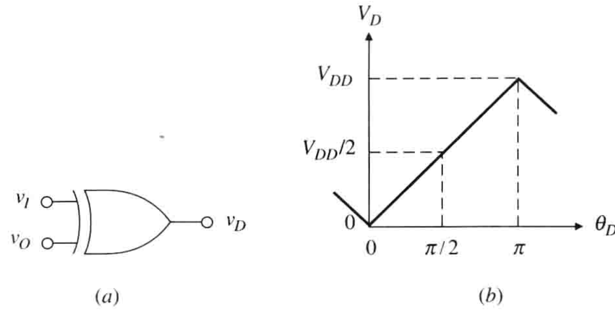


FIGURE 13.27
Type I phase comparator, and its output average V_D as a function of the input phase difference.

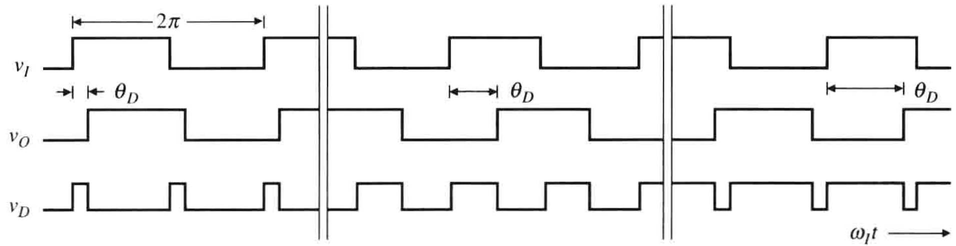


FIGURE 13.28
Typical waveforms for the Type I phase comparator in the locked condition: $\theta_D = \pi/6$ (left), $\theta_D = \pi/2$ (center), and $\theta_D = (5/6)\pi$ (right).

overdrive the multiplier and thus render the sensitivity K_d independent of the amplitude of v_I.[10]

The Type I comparator requires that both inputs have 50% duty cycles; if at least one input is asymmetrical (see Problem 13.23), the characteristic will generally be clipped, reducing the lock range. Another notorious feature of the Type I comparator is that it may allow the PLL to lock on harmonics of the input signal. Note that if v_I is absent, v_D oscillates at the same frequency as v_O, so the average of v_D is $V_D = 0.5V_{DD}$, and $\omega_O = \omega_0$.

The Type III Phase Comparator

The Type III comparator, shown in simplified form in Fig. 13.29*a*, overcomes both of the above limitations by using an edge-triggered set-reset (SR) flip-flop. As depicted in Fig. 13.30, v_D now responds only to the rising edges of v_I and v_O, regardless of the duty cycles. It is readily seen that the phase range of PC_3 is twice as large as that of PC_1, so the characteristic is as in Fig. 13.29*b*, and $K_d = V_{DD}/(2\pi) = 0.796$ V/rad.

The advantages of edge-triggering operation come at the price of higher sensitivity to noise. An input noise spike may falsely toggle the flip-flop and cause unacceptable output errors. By contrast, with a Type I comparator, an input spike is merely transmitted to the output, where it is suppressed by the loop filter.

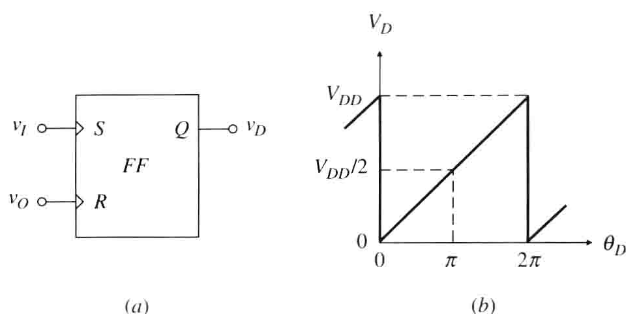


FIGURE 13.29
Type III phase comparator, and its output average V_D as a function of the input phase difference.

We observe that in the locked condition the output frequency is $\omega_D = 2\omega_I$ for PC_1, and $\omega_D = \omega_I$ for PC_3, so the ripple at the output of the loop filter is generally higher with PC_3 than with PC_1. Note that with v_I absent, PC_3 will drive ω_O as low as it can.

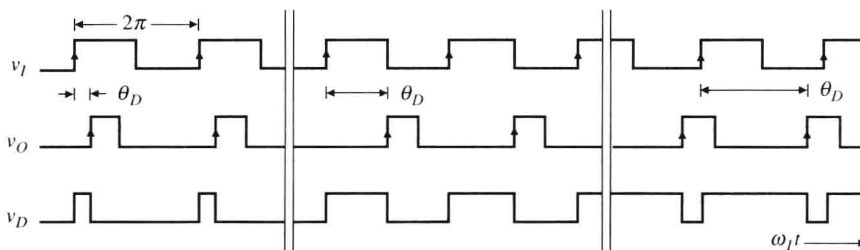


FIGURE 13.30
Typical waveforms for the Type III phase comparator in the locked condition: $\theta_D = \pi/4$ (left), $\theta_D = \pi$ (center), and $\theta_D = (7/4)\pi$ (right).

The Type II Phase Comparator

The Type II comparator differs from PC_1 and PC_3 because its output depends not only on the phase error $\theta_I - \theta_O$, but also on the frequency error $\omega_I - \omega_O$ when the loop has not yet acquired lock. Also called a *phase-frequency detector* (PFD), the circuit is shown in simplified form in Fig. 13.31*a*.

With reference to Fig. 13.32, we observe that PC_3 produces UP pulses when the rising edge of v_I leads that of v_O, DN pulses when the rising edge of v_I lags that of v_O, and no pulses when the leading edges are aligned. An UP pulse closes the MOSFET switch M_p and causes the filter capacitance C to charge toward V_{DD} via the series $R_1 + R_2$. A DN pulse closes switch M_n and discharges C toward $V_{SS} = 0$ V. Between pulses, both M_p and M_n are off, providing a high-impedance state to the filter. When PC_2 is in this state, C acts as an analog memory, retaining whatever charge it had accumulated at the end of the last UP or DN pulse.

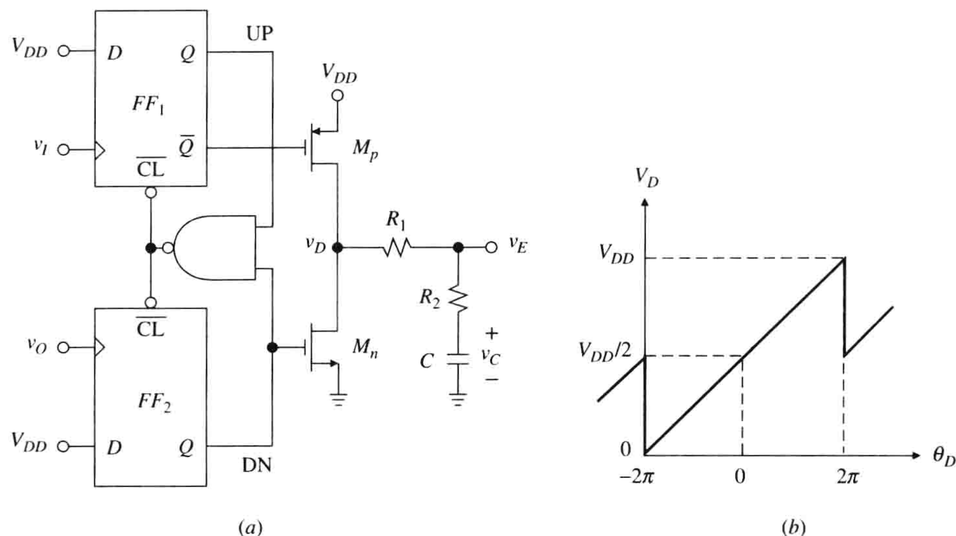


FIGURE 13.31
Type II phase comparator, and its output average V_D as a function of the input phase difference.

Clearly, we now have $v_D = v_E = v_C$. The characteristic is as in Fig. 13.31*b*, with $K_d = V_{DD}/(4\pi) = 0.398$ V/rad. For obvious reasons, PC_2 is also called a *charge-pump phase comparator.*

To appreciate its operation, suppose we initially have $\omega_I > \omega_O$. Since v_I generates more rising edges per unit time than v_O, UP will be high most of the time, pumping charge into C and thus raising ω_O. Conversely, when $\omega_I < \omega_O$, DN is high most of the time, pumping charge out of C and lowering ω_O. In either case, PC_2 will keep pumping charge until the inputs become equal *both* in frequency and in phase, or $\omega_O = \omega_I$ and $\theta_O = \theta_I$. We conclude that PC_2 approaches *ideal-integrator* behavior.

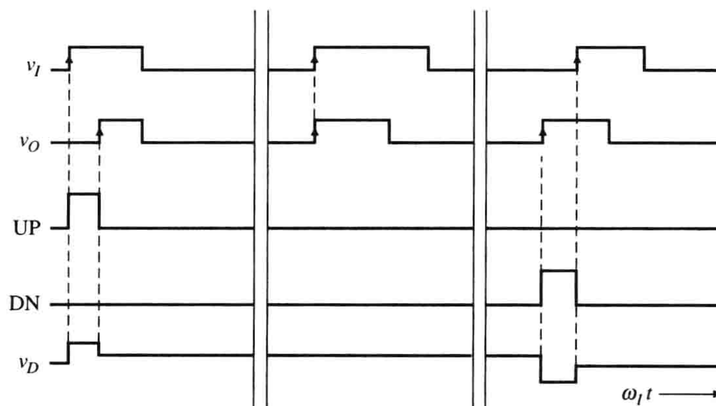


FIGURE 13.32
Typical waveforms for the Type II phase comparator for $\omega_O = \omega_I$: v_O lags v_I (left), v_O is in phase with v_I (center), and v_O leads v_I (right).

It is apparent that a PLL with a Type II comparator will lock under any condition, and that it drives the input-phase error to zero over the full frequency range of the VCO. Moreover, since the UP and DN pulses disappear entirely once the loop is locked, v_E will exhibit no ripple, so there are no unwanted phase modulation effects. The main drawback of PC_2 is its susceptibility to noise spikes, just like PC_3. Even so, PC_2 is the most popular of the three PCs. Note that with v_I absent, PC_2 will drive ω_O as low as it can.

Designing with PLLs

The design process of a PLL-based system involves a number of decisions[12] dictated by the performance specifications of the given application, along with considerations of circuit simplicity and cost. For 4046 PLLs, this process requires (*a*) the specification of the VCO parameters f_0 and $2f_R$, the choice of (*b*) the phase-detector type and (*c*) the filter type, and (*d*) the specification of the filter parameters ω_p and ω_z.

To simplify the process, computer programs are available that accept specifications by the user and translate them into actual resistance and capacitance values to meet the VCO and filter requirements. An example is the *HCMOS Phase-Locked Loop Program,* by Philips Semiconductors, which also provides important data about the loop dynamics and displays the frequency response via Bode plots. Once a PLL system has been designed, it can be simulated by computer,[7,12] for instance, using suitable SPICE macromodels.[14] However, the designer still needs a sound understanding of PLL theory to judge the results of any simulation!

Popular PLL applications[10] include FM, PM, AM, and FSK modulation/demodulation, frequency synchronization and synthesis, clock reconstruction, and motor speed control. Here we discuss two examples, FM demodulation and frequency synthesis. Other examples can be found in the end-of-chapter problems.

EXAMPLE 13.7. An FM signal is being modulated over the range of 1 MHz ± 10 kHz with a modulating frequency of 1 kHz. Using a 4046A PLL, design a circuit to demodulate such a signal.

Solution. For the VCO we let $f_0 = 1$ MHz, and choose $2f_R$ wide enough to accommodate parameter spread. Thus, let $2f_R = 0.5$ MHz. This gives $K_o = 2\pi \times 0.5 \times 10^6/2.8 = 1.122 \times 10^6$ (rad/s)/V. Using the data sheets or the aforementioned PLL program, we find that a suitable set of VCO components is $R_1 = 95.3$ kΩ, $R_2 = 130$ kΩ, and $C = 100$ pF.

Next, anticipating a noisy input signal, we choose PC_1, so $K_d = 5/\pi$ V/rad and $K_v = K_d K_o = 1.786 \times 10^6$ s^{-1}. To allow for the possibility of a weak input, we take advantage of the fact that the detector input buffers are self-biased near $V_{DD}/2$, where gain is maximized. Consequently, the input signal is ac coupled, as shown in Fig. 13.33.

Finally, to minimize cost, we use a passive lag-lead filter. Impose $\zeta = 0.707$ and choose $f_{-3\text{ dB}} > f_m$, say, $f_{-3\text{ dB}} = 10$ kHz. Proceeding as in Example 13.6, we find $\omega_p = 553$ rad/s and $\omega_z = 22.5$ krad/s, which can be met with the filter components shown in the figure.

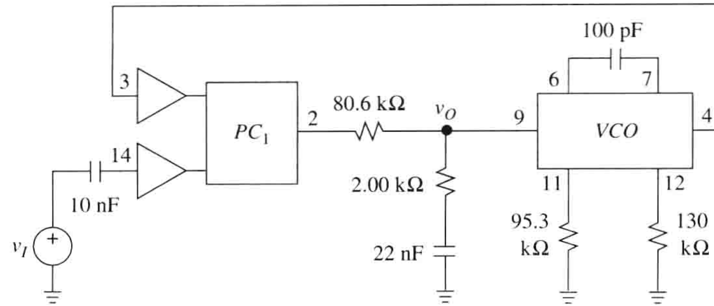


FIGURE 13.33
FM demodulator using the 4046A PLL.

Just as inserting a voltage divider within the feedback loop of an op amp increases the output voltage swing, inserting a frequency divider inside the PLL loop downstream of the VCO increases the VCO frequency. A frequency divider is implemented with a counter, and the VCO output frequency becomes $\omega_O = N\omega_I$, where N is the counter modulus. Making the counter programmable allows the synthesis of variable frequencies that are integral multiples of ω_I.

The PLL formalism still holds, but with K_o replaced by K_o/N. We observe that varying N varies also the gain factor K_v, so care must be exercised to ensure that stability and dynamics are maintained over the full range of values of N.

EXAMPLE 13.8. Using a 4046A PLL, design a circuit that accepts a 1-kHz reference frequency and synthesizes all frequencies between 1 MHz and 2 MHz in 1-kHz steps.

Solution. To span the given range, we need a programmable counter between $N_{\min} = 10^6/10^3 = 1000$ and $N_{\max} = 2000$. Choose, for instance, a 4059 counter, which allows for N to be programmed anywhere from 3 to 15,999 via a set of inputs referred to as the *jam inputs J* in the data sheets.

For the VCO, specify f_0 halfway between the extremes, or $f_0 = 1.5$ MHz, and again choose $2f_R$ wide enough, say, $2f_R = 1.5$ MHz. This gives $K_o = 3.366 \times 10^6$ (rad/s)/V. Using the data sheets or the aforementioned PLL program, we find the VCO component values $R_1 = 28.0$ kΩ, $R_2 = 287$ kΩ, and $C = 110$ pF.

Anticipating relatively clean on-board signals, we choose PC_2, so $K_d = 5/(4\pi)$ V/rad. Since N is variable, a reasonable approach[12] is to design for the geometric mean of the extremes, or for $N_{\text{mean}} = \sqrt{N_{\min}N_{\max}} = 1414$. The corresponding gain factor is then $K_{v(\text{mean})} = K_d K_o/N_{\text{mean}} = 947\ \text{s}^{-1}$.

We again use a passive lag-lead filter. Imposing $\zeta = 0.707$ and arbitrarily choosing $\omega_n = \omega_I/20 = 2\pi 10^3/20 = \pi 100$ rad/s, we obtain $\omega_p = 104$ rad/s and $\omega_z = 290$ rad/s. These parameters are met with the filter components shown in Fig. 13.34, where the wiring details of the 4059 counter have been omitted for simplicity.

Using Eq. (13.46*b*), we find $\zeta = 0.78$ for $N = 1000$, and $\zeta = 0.65$ for $N = 2000$, both of which are fairly reasonable values.

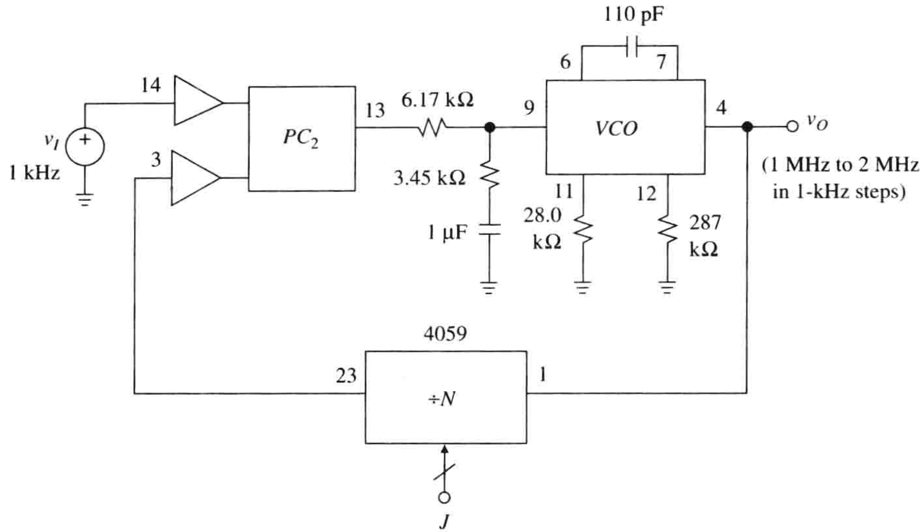


FIGURE 13.34
Frequency synthesizer using the 4046A PLL.

PROBLEMS

13.1 Log/antilog amplifiers

13.1 In the transdiode of Fig. 13.2*a* let $R = 10\ \text{k}\Omega$, $C_n + C_\mu = 20$ pF, $V_A = 100$ V, $r_d = 2\ \text{M}\Omega$, and $f_t = 1$ MHz. If $R_E = 4.3\ \text{k}\Omega$ and $C_f = 100$ pF, calculate $1/\beta_0$, $1/\beta_\infty$, f_z, and f_p for $v_I = 1$ mV, 10 mV, and 10 V; hence, confirm the linearized plots of Fig. 13.3.

13.2 Find the phase margin of the circuit of Example 13.1.

13.3 Modify the circuit of Fig. 13.4 to yield $v_O = -(2\ \text{V/dec})\log_{10}[v_I/(1\ \text{V})]$.

13.4 (*a*) Derive Eqs. (13.13) and (13.14). (*b*) Design a circuit that accepts an input voltage $-5\ \text{V} \le v_I \le +5\ \text{V}$, and gives $i_O = (10\ \mu\text{A})2^{-v_I/(1\ \text{V})}$; this circuit is useful in electronic music. (*c*) Modify the above circuit so that it gives the same output range, but for $0\ \text{V} \le v_I \le 10$ V.

13.5 The log conformity error at the upper end of the current range is due primarily to the bulk resistance of the emitter region, which can be modeled with a small resistance r_s in series with the emitter itself. (*a*) Recompute the transfer characteristic of the transdiode of Fig. 13.1*b*, but with r_s in place. If $r_s = 1\ \Omega$, what is the log conformity error at $i_I = 1$ mA? At $i_I = 0.1$ mA? (*b*) The effect of r_s can be compensated by feeding a small portion of v_I to the base of the BJT. This is achieved by lifting the base off ground, returning it to ground via a resistance R_x, and connecting a second resistance R_y between the source v_I and the base of the BJT. Sketch the modified transdiode, and show that choosing $R_y/R_x = R/r_s - 1$ will eliminate the error due to r_s.

13.6 In the log amp of Fig. 13.4 the bulk-resistance error (see Problem 13.5) can be compensated by connecting a suitable network between the base of Q_2 and the output of OA_2. Such a network consists of a resistance R_c in series with a diode D_c (cathode at the output of OA_2). Likewise, in the antilog amp of Fig. 13.6 the compensation network is connected between the base of Q_1 and the output of OA_1 (cathode at the output of OA_1). Show that the error is nulled when $R_c = (R_1 \parallel R_2)(2.2\ \text{k}\Omega)/r_s$. What is the required R_c, given that the LM394 has $r_s = 0.5\ \Omega$?

13.2 Analog multipliers

13.7 A popular multiplier application is *frequency doubling*. One way of configuring the AD534 for this operation is as follows:[6] connect X_2 and Y_1 to ground, connect X_1 and Y_2 together and drive them with a source $v_I = 10\cos\omega t$ V, connect the Out pin to Z_1 via a 10-kΩ resistor, connect Z_1 to Z_2 via another 10-kΩ resistor, and drive Z_2 with a 10-V reference voltage. (*a*) Sketch the circuit; then using the identity $\cos^2\alpha = (1 + \cos 2\alpha)/2$, obtain an expression for the output v_O. (*b*) Assuming well-regulated ±15-V supplies, design a circuit to generate the 10-V reference for Z_2.

13.8 The AD534 multiplier can be made to approximate the *sine function* within 0.5% of full scale as follows:[6] connect Y_2 to ground, connect Y_1 and Z_2 together and drive them with a source v_I, connect Y_1 to X_2 via a 10-kΩ resistor, connect X_2 to ground via an 18-kΩ resistor, connect the Out pin to Z_1 via a 4.7-kΩ resistor, connect Z_1 to X_1 via a 4.3-kΩ resistor, and connect X_1 to ground via a 3-kΩ resistor. (*a*) Sketch the circuit, derive an expression for the output v_O as a function of v_I, and calculate v_O at some significant points to verify that the circuit approximates the function $v_O = 10\sin[(v_I/10)90^\circ]$ V. (*b*) Using additional components as needed, design a circuit that accepts a triangular wave with peak values of ±5 V and gives a sine wave with the same frequency and peak values as the input.

13.9 The AD534 multiplier can be configured to yield the percentage deviation between two signals v_1 and v_2 as follows: connect X_1 and Z_1 together and drive them with v_1, connect X_2 and Y_1 to ground, drive Z_2 with v_2, connect the Out pin to Y_2 via a resistance R_1, and connect Y_2 to ground via a resistance R_2. Develop an expression for the output v_O, and specify R_1 and R_2 for $v_O = 100(v_2 - v_1)/v_1$.

13.10 Figure P13.10 shows a transducer-response linearization technique using a four-quadrant multiplier. Derive an expression for V_O as a function of δ and verify that it is linearly proportional to δ in spite of the fact that the voltage across the transducer is a nonlinear function of δ.

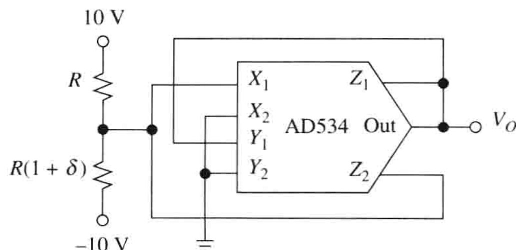


FIGURE P13.10

13.11 Using the AD534 as a voltage-controlled attenuator, design a programmable first-order low-pass filter with a dc gain of 20 dB and $f_0 = kV_C$, $0.1\text{ V} \le V_C \le 10\text{ V}$ and $k = 100$ Hz/V. *Hint:* See Problem 12.11.

13.3 Operational transconductance amplifiers

13.12 Find the transfer function of the g_m-C filter of Fig. P13.12.

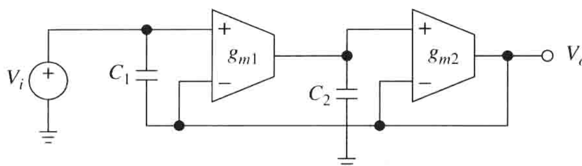


FIGURE P13.12

13.13 Design an exponential VCA such that $A = 2^{-V_C/(1\text{ V})}$ V/V, $0 \le V_C \le 10$ V. Hence, outline its calibration procedure.

13.14 Design a programmable state-variable filter with $Q = 10$, $H_{0\text{BP}} = 1$, and f_0 variable over the audio range by means of control voltage V_C as $f_0 = (20\text{ kHz})2^{-V_C/(1\text{ V})}$, $0 \le V_C \le 10$ V.

13.15 The VCA610 is a wideband VCA that accepts two signal inputs v_P and v_N and a control input V_C, and gives $v_O = A(v_P - v_N)$, where $A = 0.01 \times 10^{-V_C/(0.5\text{ V})}$ V/V for $-2\text{ V} \le V_C \le 0$. Using a VCA610 and an OPA620 wideband precision op amp, design a first-order low-pass filter with unity dc gain and programmable cutoff frequency from 100 Hz to 1 MHz.

13.16 (*a*) Sketch and label all relevant waveforms in the CCO of Fig. 13.20, and derive Eq. (13.32). (*b*) Find C so that $f_0 = 100$ kHz for $I_C = 1$ mA; next, using this CCO as basis, design a VCO such that $f_0 = (100\text{ kHz})10^{-V_C/(2\text{ V})}$, $0 \le V_C \le 10$ V. Outline its calibration procedure.

13.4 Phase-locked loops

13.17 Find the phase response $\theta_d(t)$, in degrees, in parts (*a*) and (*b*) of Example 13.3.

13.18 If we let $R_2 = 0$ in Fig. 13.23*a*, the zero is moved to infinity, resulting in a *passive lag filter*. Such a filter finds limited use because it does not allow for ω_x to be specified independently of K_v. (*a*) Verify that if we let $R_2 = 0$ in the filter of Example 13.4, the phase margin is inadequate. (*b*) Specify a new set of values for R_1 and C to ensure $\phi_m \cong 45°$ with $R_2 = 0$. What is the corresponding value of ω_x?

13.19 Repeat Example 13.4, but using an active PI filter.

13.20 Prove Eqs. (13.46) through (13.47).

13.21 A PLL has $\omega_0 = 2\pi 10^6$ rad/s, $K_d = 0.2$ V/rad, $K_a = 1$ V/V, and $K_o = \pi 10^6$ (rad/s)/V. Design an active PI filter for a loop time constant of approximately 100 periods of the free-running frequency and $Q = 0.5$.

13.22 If a 0.1-μF capacitance is connected in parallel with R_2 in the loop filter of Example 13.6, find how it affects ω_x and ϕ_m.

13.5 Monolithic PLLs

13.23 (*a*) Sketch and label the average V_D versus θ_D for a Type I phase comparator if the duty cycles of v_I and v_O are $D_I = \frac{1}{2}$ and $D_O = \frac{1}{3}$. (*b*) Repeat, but with $D_I = \frac{1}{3}$ and $D_O = \frac{1}{2}$. Comment.

13.24 Sketch v_I, v_O, UP, DN, and v_D for a Type II detector if (*a*) ω_I is slightly higher than ω_O, (*b*) ω_I is slightly lower than ω_O, (*c*) $\omega_I \gg \omega_O$, and (*d*) $\omega_I \ll \omega_O$.

13.25 A certain CMOS PLL is powered between 5 V and 0 V, and uses a Type I phase comparator and a VCO with $K_o = 5$ MHz/V and $f_0 = 10$ MHz for $v_E = 2.5$ V. (*a*) Design a passive lag-lead filter for $\omega_n = 2\pi 5$ krad/s and $Q = 0.5$. (*b*) Sketch v_I, v_O, v_D, and v_E for the case in which the loop is locked to an input frequency of 7.5 MHz.

13.26 Find $v_e(t)$ in the FM demodulator of Example 13.7.

13.27 Dual-slope ADCs are clocked at a frequency that is locked to the ac line frequency f_{line} in order to reject line-induced noise. Using a 4046A PLL, design a circuit that accepts f_{line} (either 60 Hz or 50 Hz) and gives $f_{\text{CK}} = 2^{16} \times f_{\text{line}}$. Specify as many parameters and components as you can in your circuit.

13.28 Using a 4046A for phase detection and an 8038 as VCO, design a circuit that generates a 1-kHz sine wave synchronized on a 1-MHz crystal oscillator.

13.29 An FSK signal v_I alternates between $f_L = 1200$ Hz (logic 0) and $f_H = 2400$ Hz (logic 1). One way[13] to decode this signal with a 4046A PLL is to use PC_3, a loop filter consisting of a plain RC stage with $1/(2\pi RC) = f_H$, the VCO with $f_0 = (f_L + f_H)/2 = 1.8$ kHz and $2f_R = 2$ kHz, and a positive edge-triggered latch flip-flop of the type of Fig. 13.31, with v_I as the D input and the VCO output v_O as the clock; the $\overline{Q}$ output of the flip-flop is the FSK decoder output. Draw the circuit; then sketch and label v_I, the average of v_E, v_O, and $\overline{Q}$ both for $f_I = f_L$ and $f_I = f_H$. What is the distinguishing feature of PC_3 that makes it attractive in this application?

REFERENCES

1. D. H. Sheingold, ed., *Nonlinear Circuits Handbook,* Analog Devices, Norwood, MA, 1974.
2. P. R. Gray, P. J. Hurts, S. H. Lewis, and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits,* 5th ed., John Wiley & Sons, New York, 2009. ISBN 978-0-470-24599-6.
3. "Theory and Applications of Logarithmic Amplifiers," Texas Instruments Application Note AN-311, http://www.ti.com/lit/an/snoa575b/snoa575b.pdf.
4. C. Kitchin and L. Counts, *RMS to DC Conversion Application Guide,* 2d ed., http://www.analog.com/static/imported-files/design_handbooks/17505283758302176206322RMS-toDC_Cover-Section-I.pdf.
5. B. Gilbert, "Translinear Circuits—25 Years On," *Electronic Engineering:* Part I, August 1993, pp. 21–24; Part II, September 1993, pp. 51–53; Part III, October 1993, pp. 51–56.

6. D. H. Sheingold, ed., *Multiplier Application Guide,* Analog Devices, Norwood, MA, 1978.
7. D. A. Johns and K. W. Martin, *Analog Integrated Circuit Design,* John Wiley & Sons, New York, 1997.
8. C. Toumazou, F. J. Lidgey, and D. G. Haigh, eds., *Analogue IC Design: The Current-Mode Approach,* IEEE Circuits and Systems Series, Peter Peregrinus Ltd., London, U.K., 1990.
9. R. Schaumann, M. S. Ghausi, and K. R. Laker, *Design of Analog Filters: Passive, Active RC, and Switched Capacitor,* Prentice Hall, Englewood Cliffs, NJ, 1990.
10. A. B. Grebene, *Bipolar and MOS Analog Integrated Circuit Design,* John Wiley & Sons, New York, 1984.
11. F. M. Gardner, *Phaselock Techniques,* 2d ed., John Wiley & Sons, New York, 1979.
12. R. E. Best, *Phase-Locked Loops: Theory, Design, and Applications,* 3d ed., McGraw-Hill, New York, 1997.
13. *CMOS Phase-Locked Loops,* Philips Semiconductors, Sunnyvale, CA, June 1995.
14. J. A. Connelly and P. Choi, *Macromodeling with SPICE,* Prentice Hall, Englewood Cliffs, NJ, 1992.

INDEX

G

H

M

N

T

U

V

推荐阅读



电路基础（英文版·第5版）

作者：（美）Charles K. Alexander 等 于歆杰 注释 ISBN：978-7-111-41184-0 定价：129.00元

本书是电类各专业“电路”课程的一本经典教材，被美国众多名校采用，是美国最有影响力的“电路”课程教材之一。本书每章开始增加了中文“导读”，适合用做高校“电路”课程双语授课或英文授课的教材。本书前4版获得了极大的成功，第5版以更清晰、更容易理解的方式阐述了电路的基础知识和电路分析方法，并反映了电路领域的最新技术进展。全书总共包括2447道例题和各类习题，并在书后给出了部分习题答案。

交直流电路基础：系统方法

作者：（美）Thomas L. Floyd 译者：殷瑞祥 等 ISBN：978-7-111-45360-4 定价：99.00元

本书是知名作者Folyd的最新力作，在国外被广泛使用。本书系统介绍了直流和交流电路理论，强调直流/交流电路基本概念在实际系统中的应用。全书丰富的实例，有助于学生的理解系统模块、接口和输入/输出信号之间的关系。书中实例使用Multisim进行仿真，并提出在模拟电路与系统和排除故障中存在的问题及解决方法。本书可作为电子信息、电气工程、自动化等电类专业的电路课程教材。

应用电路分析（英文版）

作者：（美）Matthew N. O. Sadiku 等 ISBN：978-7-111-41781-1 定价：89.00元

本书可作为高等院校电类专业“电路分析”双语课的教材，以更清晰、生动、易于理解的方式来阐述电路分析的方法。全书分为两部分，第一部分包括第1~10章，主要介绍直流电路；第二部分包括第11~19章，主要介绍交流电路。本书可以作为大学两学期或三学期的教材，授课教师也可选择适当的章节，将其用作一学期课程的教材。